

## **La interacción entre factores en el análisis de varianza: errores de interpretación**

La interacción entre factores en el análisis de varianza es interpretada erróneamente en la investigación empírica con mucha frecuencia. En una revisión de 272 artículos se ha encontrado una interpretación basada en el análisis de los efectos simples en el 72,9% de las interacciones de primer orden y el 35,7% de las interacciones de segundo orden. El 18,6% de las interacciones de primer orden y el 60% de las interacciones de segundo orden no se analizaron o no se interpretaron. El 8,5% y 4,3% respectivamente, fueron analizadas e interpretadas correctamente. Los programas informáticos utilizados habitualmente no permiten realizar directamente las comparaciones necesarias para evaluar una interacción significativa en los diseños factoriales con grupos al azar. Para contribuir a erradicar este problema se muestra como diseñar algunas de las comparaciones lineales que permiten aislar el efecto de la interacción y se explica detalladamente como realizar esas comparaciones con el SPSS.

Palabras clave: análisis de varianza, interacción, ANOVA

*Interaction in ANOVA: Misconceptions.* In empirical research, misconceptions related to the interpretation of interaction in analysis of variance are frequent. After review of 272 published papers, interpretations based on simple main effects were found in 72.9% of those analyzing two-way interactions and in 35.7% of those analyzing three-way interactions. Interaction were not analyzed or interpreted in 18.6% of two-way interactions and 60% of three-way interactions, and only 8.5% and 4.3%, respectively, were correctly analyzed and interpreted. Most of the statistical packages do not allow performing directly the comparisons needed to assess a significant interaction in factorial designs with random groups. In order to solve this problem, we herein show how to design some linear comparisons needed to isolate interaction effects, and how to compute them with SPSS is explained in detail.

Key words: analysis of variance, interaction, ANOVA



**Facultad de Psicología**

# **La interacción entre factores en el análisis de varianza: errores de interpretación**

Tesis doctoral

**Autor: Jesús Garrido García**

Director: Antonio Pardo Merino

Madrid, 2008



### Agradecimientos:

A Julia, que nunca leerá este trabajo, pero sin cuyo soporte no se habría gestado.

A mi madre, siempre dispuesta a disculpar los ratos que este trabajo le ha robado.

A mi director, de quien he aprendido tantas y tantas cosas y del que espero aprender aún muchas más.

A todos los miembros del área de Metodología de las Ciencias del Comportamiento de la Facultad de Psicología de la UAM, que me han acogido como a uno más y cuyo apoyo me estimula, rejuvenece y me hace mejor de lo que soy.



# Índice

## 1. INTRODUCCIÓN

1.1. Objetivos del trabajo.....	11
---------------------------------	----

## 2. EL CONCEPTO DE INTERACCIÓN

2.1. Interacción de primer orden.....	13
2.2. Interacción de segundo orden.....	17

## 3. ERRORES COMETIDOS AL ANALIZAR Y/O INTERPRETAR LA INTERACCIÓN

3.1. Errores cometidos al analizar y/o interpretar la interacción de primer orden.....	24
A) Error consistente en no analizar o no interpretar la interacción de primer orden.....	27
B) La interacción de primer orden se interpreta recurriendo al análisis de los efectos simples.....	29
C) Interpretación correcta de la interacción de primer orden.....	36
3.2. Errores cometidos al analizar y/o interpretar la interacción de segundo orden.....	37
A) Error consistente en no analizar o no interpretar la interacción de segundo orden.....	39
B) La interacción de primer orden se interpreta recurriendo al análisis de los efectos simples.....	40
C) Interpretación correcta de la interacción de segundo orden.....	45
3.3. Por qué se recurre al análisis de los efectos simples para interpretar una interacción significativa.....	46
3.3.1. Discrepancias entre manuales.....	46
3.3.2. Limitaciones de los programas informáticos.....	54

#### **4. EJEMPLOS DE ERRORES QUE SE COMENTEN AL ANALIZAR LA INTERACCIÓN**

4.1. Ejemplos de errores que se comenten al analizar la interacción de primer orden.....	57
A) Ejemplos de errores consistentes en no analizar y/o no interpretar la interacción de primer orden.....	57
B) Ejemplos de errores en los que la interacción de primer orden se interpreta recurriendo al análisis de los efectos simples.....	77
C) Ejemplos de interpretación correcta de la interacción de primer orden.....	109
4.2. Ejemplos de errores que se comenten al analizar la interacción de segundo orden.....	117
A) Ejemplos de errores consistentes en no analizar y/o no interpretar la interacción de segundo orden.....	117
B) Ejemplos de errores en los que la interacción de segundo orden se interpreta recurriendo al análisis de los efectos simples.....	119
B.1) Interpretación a partir de los efectos simples simples .....	120
B.2) Interpretación a partir de los efectos simples .....	129
B.3) Interpretación a partir de los efectos simples de interacción ...	134
C) Ejemplos de interpretación correcta de la interacción de segundo orden.....	140

#### **5. CÓMO EFECTUAR CONTRASTES PARA INTERPRETAR LA INTERACCIÓN**

5.1. El caso de dos factores.....	146
5.1.1. Diseños 2×3 .....	147
5.1.2. Diseños 3×3 .....	152
5.1.3. Diseños 3×4 .....	155
5.1.4. El procedimiento ANOVA de un factor .....	160
5.1.5. La sentencia LMATRIX.....	170
5.2. El caso de tres factores.....	174
5.2.1. Diseños 2×2×3 .....	178
5.2.2. Diseños 2×3×3 .....	182
5.2.3. Diseños 3×3×3 .....	189
5.2.4. El procedimiento ANOVA de un factor .....	193
5.2.5. La sentencia LMATRIX .....	204

<b>6. CONCLUSIONES.....</b>	<b>207</b>
<b>7. REFERENCIAS BIBLIOGRÁFICAS.....</b>	<b>209</b>
<b>8. ANEXOS</b>	
Anexo 1: Artículos revisados .....	215
Anexo 2: Valoración de los artículos revisados.....	231
Anexo 3: Índice de tablas.....	238
Anexo 4: Índice de figuras.....	240





# 1. Introducción

Los modelos factoriales de análisis de varianza son ampliamente utilizados como estrategia de análisis de datos en muy diversas áreas de conocimiento. Entre las ventajas que justifican el uso tan extendido de estos modelos, quizá la más destacable sea que ofrecen la posibilidad de estudiar el efecto de la interacción entre factores.

El concepto de interacción está explícita y abundantemente tratado en la literatura estadística y desde el punto de vista teórico, tiene un significado inequívoco (ver, por ejemplo, Hays, 1994, págs. 479-480; Jaccard, 1998, págs. 3-6; Keppel y Wickens, 2004, págs. 198-206; Maxwell y Delaney, 2004, págs. 277-280; Winer, Brown y Michels, 1991, págs. 296-298; etc.). Sin embargo, cuando se trata, no de *definir* la interacción, sino de *analizarla e interpretarla*, no parece que las cosas estén tan claras.

Ya en 1970, Marascuilo y Levin alertaron sobre lo que dieron en llamar el error de tipo IV y que definieron como la interpretación incorrecta del rechazo correcto de una hipótesis nula. Al definir este tipo de error se refirieron explícitamente a las estrategias utilizadas para interpretar una interacción significativa en el contexto del análisis de varianza. Levin y Marascuilo (1972), señalaron que en los estudios aplicados es frecuente encontrar que las interacciones significativas son frecuentemente analizadas y explicadas de forma incorrecta. Rosnow y Rosenthal (1989b) llegaron a afirmar (pág. 1.282) que la interacción es “probably the universally most misinterpreted empirical result in psychology”. Años después, la interacción entre factores seguía siendo un resultado frecuentemente mal interpretado. En 1993, Zuckerman, Hodgins, Zuckerman,

y Rosenthal, publicaron una encuesta, contestada en un 89% por profesores universitarios, en la que una de las preguntas era la siguiente:

One way to understand interaction effects is to calculate the simple effects, i.e. the effect of each factor at a constant level of the other factor. In a  $2 \times 2$  design, for example (see table below), an interaction effect can be interpreted by conducting two  $t$  test; one  $t$  comparing cell *a* to cell *b* and another  $t$  comparing cell *c* to cell *d*. It is a correct approach to the interpretation of interactions?

<i>a</i>	<i>b</i>
<i>c</i>	<i>d</i>

\_\_\_ It depends \_\_\_ Yes \_\_\_ No \_\_\_ Blank

para la que la respuesta correcta era “No”. Las respuestas recibidas fueron las siguientes: “Depende”, 112 encuestados; “Sí”, 88 encuestados; “No”, 353 encuestados; En blanco, 8 encuestados. Las respuestas “Depende” y “Sí” frecuentemente eran acompañadas con comentarios del estilo “sólo si la interacción es significativa”

En 1991, Ottenbacher realizó una revisión de los artículos publicados en las revistas *New England Journal of Medicine*, *Pediatrics* y *Journal of the American Medical Association* entre los años 1982 y 1989, identificando los artículos en los que se utilizaba un ANOVA de dos factores con una interacción significativa. En total seleccionó 83 artículos que clasificó en las categorías, correctamente interpretados, interpretación basada en la comparación de las medias de las casillas, interpretación de la interacción en términos de efectos principales o no interpretación en absoluto. En el apartado de resultados comenta:

... 20 of the studies or approximately twenty-four per cent provided an analysis and explanation of a statistically significant interaction that correspond to the correct interpretation based on a two-way ANOVA... Forty-six per cent [of total 83 studies], or 38 of 55 studies that contained an incorrect interpretation were

classified as category 2 [cell means interpretation]. The interpretation of significant interaction in these studies was as a comparison between cell means within one factor. This simple main effect interpretation does not correspond to the interaction model of ANOVA.

En la actualidad la interacción entre factores sigue siendo un resultado frecuentemente mal interpretado, como tendremos ocasión de comprobar. Según veremos: (1) en los manuales de diseño y análisis de datos y en los artículos que discuten de forma explícita el concepto de interacción, no parece existir un acuerdo generalizado sobre la estrategia que debe seguirse para interpretar correctamente una interacción significativa; y (2) la mayoría de los trabajos de investigación revisados contienen errores de distinta naturaleza tanto en la forma de analizar la interacción como, sobre todo, en la de interpretarla.

## 1.1. Objetivos del trabajo

Este trabajo tiene el doble objetivo de:

1. Identificar el tipo de errores que se cometen cuando se analiza e interpreta el efecto de la interacción entre factores en el análisis de varianza (tanto en lo referente a las interacciones de primer como de segundo orden).
2. Proponer una estrategia analítica que ayude a los investigadores en psicología y ciencias afines a analizar e interpretar correctamente el efecto de la interacción entre factores.

Para cubrir el primer objetivo se han revisado los artículos publicados durante los años 2000 a 2006 en cuatro revistas de psicología editadas en España de las de mayor factor

de impacto: la *Revista Internacional de Psicología Clínica y de la Salud*, *Psicothema*, *Psicológica* y *The Spanish Journal of Psychology*. Se han definido los tipos de errores encontrados y se ha clasificado cada estudio en función de esa tipología.

Para cubrir el segundo objetivo: 1) se ha realizado una revisión de los procedimientos que propone la literatura estadística para analizar correctamente una interacción significativa; y 2) se ha estudiado la posibilidad de aplicarlos con las herramientas informáticas de uso más frecuente entre los investigadores en psicología. Todo ello, intentando diseñar propuestas susceptibles de ser entendidas y utilizadas por los investigadores en psicología y ciencias afines.

## 2. El concepto de interacción

### 2.1. Interacción de primer orden

El concepto de interacción entre factores admite varias formulaciones, todas ellas equivalentes. Desde un punto de vista *no formal*, decimos que existe interacción entre dos factores cuando el efecto de uno de ellos sobre la variable dependiente no es el mismo en todos los niveles del otro factor. Esto equivale a afirmar que existe interacción cuando el resultado de la combinación de dos factores difiere de la suma de los efectos principales de esos factores. Revisando la bibliografía sobre el tema encontramos definiciones que se centran en ambos puntos de vista:

Everitt y Howell (2005, págs. 930-931):

Interactions as conditional effects... If there is interaction, the conditional effects of one variable differs across values of the other variable

Interactions as nonadditive effects... the combination of two or more variables does not produce an outcome that is the sum of their individual effects.

Interactions as cell residuals... By cell residuals is meant the discrepancy between the cell mean and the grand mean that would not be expected from the additive effects of each variable. When there is an interaction between variables, the cell residuals are nonzero and are a pure measures of the amount of interaction.

Kirk (1995, pág. 367):

...Two treatments are said to interact if differences in performance under the levels of one treatment are different at two or more levels of the other treatment. An interaction

can be thought of as the joint effects of treatments that are different from the sum of their individual effects.

Maxwell y Delaney (2004, pág. 277):

...To say that an interaction exists in our data means that the [1<sup>st</sup> factor] biofeedback effect in presence of [2<sup>nd</sup> factor] drug therapy is different from the average effect of [1<sup>st</sup> factor] biofeedback.

Para poder presentar una definición *formal* de la interacción, consideremos la notación propuesta en la Tabla 2.1 para la configuración de un diseño factorial 2×3, con los factores  $A$  y  $B$  (con niveles  $A_j$  y  $B_k$ ;  $j=1, 2$ ;  $k=1, 2, 3$ ), la media total ( $\mu$ ), las medias marginales ( $\mu_{j+}$  y  $\mu_{+k}$ ) y las medias de las casillas ( $\mu_{jk}$ ).

**Tabla 2.1. Notación utilizada en un diseño 2×3**

	$B_1$	$B_2$	$B_3$	
$A_1$	$\mu_{11}$	$\mu_{12}$	$\mu_{13}$	$\mu_{1+}$
$A_2$	$\mu_{21}$	$\mu_{22}$	$\mu_{23}$	$\mu_{2+}$
	$\mu_{+1}$	$\mu_{+2}$	$\mu_{+3}$	$\mu$

En el modelo de ANOVA de dos factores de efectos fijos y grupos aleatorios, el efecto de la interacción se define como (ver, por ejemplo, Winer y cols, 1991, pág. 318):

$$(\alpha\beta)_{jk} = \mu_{jk} - \mu_{j+} - \mu_{+k} + \mu \quad [2.1]$$

De acuerdo con esta definición, existe interacción cuando  $(\alpha\beta)_{jk} \neq 0$  para algún  $j$  o  $k$ ; y no existe interacción cuando  $(\alpha\beta)_{jk} = 0$  para todo  $j$  y  $k$ . Ahora bien, existen al menos dos maneras alternativas de interpretar la definición [2.1] (ver Jaccard, 1998, págs. 3-10):

1. Como la *desviación que experimentan las medias de las casillas respecto de los efectos principales de los factores*:

- 1.a. No interacción:  $\mu_{jk} = \mu_{j+} + \mu_{+k} - \mu$  (para todo  $j$  y  $k$ )

$$1.b. \text{ Interacción: } \mu_{jk} \neq \mu_{j+} + \mu_{+k} - \mu \text{ (para algún } j \text{ o } k) \quad [2.2]$$

Según esta definición, existe interacción cuando la media de una o más casillas no es función directa de sus respectivas medias marginales. Esto significa que, si existe interacción, el efecto de *la combinación* de los factores *A* y *B* difiere de *la suma* de los efectos de *A* y *B*.

2. Como *diferencias entre las medias de las casillas y las medias marginales*:

$$2.a. \text{ No interacción: } \mu_{jk} - \mu_{j'k} = \mu_{j+} - \mu_{j'+} \text{ (para todo } j, j' \text{ y } k) \quad [2.3]$$

$$2.b. \text{ Interacción: } \mu_{jk} - \mu_{j'k} \neq \mu_{j+} - \mu_{j'+} \text{ (para algún } j, j' \text{ o } k)$$

Según esta definición, existe interacción cuando la diferencia entre las medias de dos casillas de la misma columna (lo mismo vale para las filas) no es igual que la diferencia entre sus correspondientes medias marginales.

Las expresiones [2.2] y [2.3] se deducen, ambas, de [2.1]; por tanto, son equivalentes.

En efecto, según la expresión [2.2], cuando no existe interacción:

$$\mu_{11} = \mu_{1+} + \mu_{+1} - \mu \quad \text{y} \quad \mu_{21} = \mu_{2+} + \mu_{+1} - \mu$$

$$\text{de aquí se sigue: } \mu_{11} - \mu_{1+} - \mu_{+1} = -\mu \quad \text{y} \quad \mu_{21} - \mu_{2+} - \mu_{+1} = -\mu$$

$$\text{por tanto: } \mu_{11} - \mu_{1+} - \mu_{+1} = \mu_{21} - \mu_{2+} - \mu_{+1}$$

$$\text{lo cual lleva a: } \mu_{11} - \mu_{21} = \mu_{1+} - \mu_{2+}$$

que no es otra cosa que la expresión [2.3]. De este argumento se desprende que es irrelevante tomar [2.2] o [2.3] como referente para la discusión. Sin embargo, las interpretaciones basadas en [2.3] suelen resultar más fáciles de entender.



La primera formulación (basada en [2.2]) ha recibido especial atención por parte de Rosnow y Rosenthal (1989a, 1989b, 1991, 1995, 1996), quienes han llevado el argumento al extremo defendiendo repetidamente la idea de que para poder interpretar correctamente el efecto de la interacción hay que despojarlo de todos los elementos que incluye. Es así como se llega a las *medias residualizadas* o *residuos de interacción* que en un diseño de dos factores (ver ecuación 2.1) se obtienen eliminando de cada casilla el efecto de su fila y de su columna, es decir, restando a la media de cada casilla ( $\mu_{jk}$ ) sus correspondientes medias marginales ( $\mu_{j+}$  y  $\mu_{+k}$ ) y sumando la media total ( $\mu$ ). Cuando no existe efecto de la interacción estos residuos valen cero. Estos residuos, siempre según Rosnow y Rosenthal, son los únicos que informan cabalmente sobre el efecto de la interacción. Pero algunos autores son muy críticos con este enfoque (ver Meyer, 1991; Petty, Fabrigar, Wegener y Priester, 1996), argumentando entre otras cosas que las medias residualizadas pueden dar lugar a resultados ininterpretables en función de la naturaleza de las variables estudiadas. Por otra parte, este enfoque es muy poco utilizado por los investigadores. En una revisión realizada por Rosnow y Rosenthal (1989a) solo en el 1% de 191 artículos revisados se interpretó la interacción a partir de los *residuos de interacción*.

Unas sencillas transformaciones permiten comprobar que la definición [2.3] implica que, si existe interacción, la diferencia entre los niveles  $A_1$  y  $A_2$  no es la misma en los tres niveles de  $B$ ; y lo mismo vale decir de la diferencia entre los niveles  $B_1$ ,  $B_2$  y  $B_3$  en los dos niveles de  $A$ .

Este es el significado de la interacción entre los factores  $A$  y  $B$ . Y ésta es la idea que se pretende destacar: la interpretación de la interacción requiere *comparar diferencias*.

Cuando la situación se simplifica a un diseño  $2 \times 2$ , unas sencillas transformaciones permiten comprobar que la ecuación [2.3] equivale a:

$$\mu_{11} - \mu_{21} = \mu_{12} - \mu_{22} \quad [2.4]$$

La comparación propuesta en [2.4] es la que corresponde al único grado de libertad asociado a la interacción en un diseño  $2 \times 2$ . Por tanto, si el estadístico  $F$  asociado al efecto de la interacción es significativo, una interpretación basada en las diferencias comparadas en [2.4] agota el significado de la interacción, lo cual implica que no es necesario recurrir a comparaciones adicionales para interpretar una interacción significativa. Debe tenerse en cuenta que, si se verifica [2.4], también se verifican:

$$\mu_{11} - \mu_{12} = \mu_{21} - \mu_{22} \quad \text{y} \quad \mu_{11} + \mu_{22} = \mu_{12} + \mu_{21}$$

por tanto, una interacción significativa puede interpretarse recurriendo a cualquiera de estas tres comparaciones de diferencias.

## 2.2. Interacción de segundo orden

En un diseño de tres factores ( $A$ ,  $B$ ,  $C$ ) existen diversos efectos de interés. Además de los efectos principales de cada factor y el de las interacciones de primer orden entre cada par de factores (las cuales se definen exactamente de la misma manera que en un diseño de dos factores), existe el efecto de la interacción de segundo orden, es decir, de la interacción entre los tres factores.

La interacción de segundo orden se define de forma parecida a como se hace con la de primer orden. Desde un punto de vista *no formal* decimos que existe interacción de

segundo orden ( $ABC$ ) cuando la interacción de primer orden entre los dos primeros factores ( $AB$ ) no es constante en todos los niveles del tercer factor ( $C$ ). Por supuesto, lo mismo vale decir de la interacción  $AC$  respecto de los niveles de  $B$  y de la interacción  $BC$  respecto de los niveles de  $A$ : existe interacción de segundo orden cuando el efecto de cualquiera de las interacciones de primer orden no es constante (no es el mismo) en todos los niveles del tercer factor.

Para ofrecer una definición formal de la interacción de segundo orden consideremos la notación propuesta en la Tabla 2.2 para un diseño factorial  $2 \times 2 \times 2$ , con los factores  $A$ ,  $B$  y  $C$  (con niveles  $A_j$ ,  $B_k$  y  $C_l$ ;  $j = 1, 2$ ;  $k = 1, 2$ ;  $l = 1, 2$ ), la media total ( $\mu$ ), algunas medias marginales ( $\mu_{j++}$ ,  $\mu_{+k+}$ ,  $\mu_{++l}$  y  $\mu_{++l}$ ) y las medias de las casillas ( $\mu_{jkl}$ ).

**Tabla 2.2. Notación utilizada en un diseño  $2 \times 2 \times 2$**

	$B_1$		$B_2$		
	$C_1$	$C_2$	$C_1$	$C_2$	
$A_1$	$\mu_{111}$	$\mu_{112}$	$\mu_{121}$	$\mu_{122}$	$\mu_{1++}$
$A_2$	$\mu_{211}$	$\mu_{212}$	$\mu_{221}$	$\mu_{222}$	$\mu_{2++}$
	$\mu_{+11}$	$\mu_{+12}$	$\mu_{+21}$	$\mu_{+22}$	
	$\mu_{+1+}$		$\mu_{+2+}$		$\mu$

En el modelo de ANOVA de tres factores de efectos fijos y grupos aleatorios, el efecto de la interacción se define como:

$$\begin{aligned}
 (\alpha\beta\gamma)_{jkl} &= \mu_{jkl} - \mu - [\alpha_j + \beta_k + \gamma_l + (\alpha\beta)_{jk} + (\alpha\gamma)_{jl} + (\beta\gamma)_{kl}] \\
 &= \mu_{jkl} - \mu - (\mu_{j++} - \mu) - (\mu_{+k+} - \mu) - (\mu_{++l} - \mu) - (\mu_{jk+} - \mu_{j++} - \mu_{+k+} + \mu) \\
 &\quad - (\mu_{j+l} - \mu_{j++} - \mu_{++l} + \mu) - (\mu_{+kl} - \mu_{+k+} - \mu_{++l} + \mu) = \\
 &= \mu_{jkl} - (-\mu_{j++} - \mu_{+k+} - \mu_{++l} + \mu_{jk+} + \mu_{j+l} + \mu_{+kl}) - \mu
 \end{aligned} \tag{2.5}$$

De acuerdo con esta definición, existe interacción de segundo orden cuando  $(\alpha\beta\gamma)_{jkl} \neq 0$  para algún  $j, k$  o  $l$ ; y no existe interacción cuando  $(\alpha\beta\gamma)_{jkl} = 0$  para todo  $j, k$  y  $l$ . Ahora bien, existen al menos dos maneras alternativas de interpretar esta definición:

1. Como la *desviación que experimentan las medias de las casillas respecto de sus correspondientes medias marginales*:

$$\begin{aligned}
 1.a. \text{ No interacción: } & \mu_{jkl} = -\mu_{jk+} - \mu_{j+l} - \mu_{+kl} + \mu_{j++} + \mu_{+k+} + \mu_{++l} + \mu \\
 & \text{(para todo } j, k \text{ y } l) \\
 1.b. \text{ Interacción: } & \mu_{jkl} \neq -\mu_{jk+} - \mu_{j+l} - \mu_{+kl} + \mu_{j++} + \mu_{+k+} + \mu_{++l} + \mu \\
 & \text{(para algún } j, k \text{ o } l)
 \end{aligned}
 \tag{2.6}$$

2. Como *diferencias entre las interacciones de primer orden en los diferentes niveles del tercer factor*:

$$\begin{aligned}
 2.a. \text{ No interacción: } & \mu_{jkl} - \mu_{j'kl} - (\mu_{jk'l} - \mu_{j'k'l}) = \mu_{jkl'} - \mu_{j'kl'} - (\mu_{jk'l'} - \mu_{j'k'l'}) \\
 & \text{(para todo } j, j', k, k', l \text{ y } l') \\
 2.b. \text{ Interacción: } & \mu_{jkl} - \mu_{j'kl} - (\mu_{jk'l} - \mu_{j'k'l}) \neq \mu_{jkl'} - \mu_{j'kl'} - (\mu_{jk'l'} - \mu_{j'k'l'}) \\
 & \text{(para algún } j, j', k, k', l \text{ o } l')
 \end{aligned}
 \tag{2.7}$$

En esta segunda forma de entender la interacción se está enfatizando el hecho de que no existe interacción de segundo orden cuando todas las interacciones de primer orden son iguales en todos los niveles del tercer factor. Y en este contexto cobra especial importancia el concepto de *efecto simple de interacción*, el cual se define como:

$$(\alpha\beta)_{jk} \text{ en } C_l = AB|C_l = \mu_{jkl} - \mu_{j'kl} - (\mu_{jk'l} - \mu_{j'k'l})
 \tag{2.8}$$

Si no existe interacción, los efectos simples de interacción entre dos factores son iguales en todos los niveles del tercer factor. Si dos efectos simples de interacción no son iguales, entonces existe interacción de segundo orden.

Las expresiones [2.6] y [2.7] se deducen, ambas, de [2.5]; por tanto, son equivalentes. En efecto, según la expresión [2.6], cuando no existe interacción de segundo orden se verifica:

$$\mu_{111} = -\mu_{11+} - \mu_{1+1} - \mu_{+11} + \mu_{1++} + \mu_{+1+} + \mu_{++1} + \mu \quad (\text{media de } A_1B_1 \text{ en } C_1)$$

$$\mu_{211} = -\mu_{21+} - \mu_{2+1} - \mu_{+11} + \mu_{2++} + \mu_{+1+} + \mu_{++1} + \mu \quad (\text{media de } A_2B_1 \text{ en } C_1)$$

Restando ambas expresiones se obtiene la diferencia entre la media de la combinación  $A_1B_1$  y la media de la combinación  $A_2B_1$  en el primer nivel de  $C$ :

$$\mu_{111} - \mu_{211} = \mu_{21+} + \mu_{2+1} - \mu_{11+} - \mu_{1+1} + \mu_{1++} - \mu_{2++}$$

Además,

$$\mu_{121} = -\mu_{12+} - \mu_{1+1} - \mu_{+21} + \mu_{1++} + \mu_{+2+} + \mu_{++1} + \mu \quad (\text{media de } A_1B_2 \text{ en } C_1)$$

$$\mu_{221} = -\mu_{22+} - \mu_{2+1} - \mu_{+21} + \mu_{2++} + \mu_{+2+} + \mu_{++1} + \mu \quad (\text{media de } A_2B_2 \text{ en } C_1)$$

Restando ambas expresiones se obtiene la diferencia entre la media de la combinación  $A_1B_2$  y la media de la combinación  $A_2B_2$  en el primer nivel de  $C$ :

$$\mu_{121} - \mu_{221} = \mu_{22+} + \mu_{2+1} - \mu_{12+} - \mu_{1+1} + \mu_{1++} - \mu_{2++}$$

Ahora bien, la interacción entre los factores  $A$  y  $B$  en el nivel  $C_1$  (o efecto simple de interacción  $AB$  en  $C_1$ ) viene dada por:

$$AB|C_1 = \mu_{jk1} - \mu_{j'k1} - (\mu_{jk'1} - \mu_{j'k'1})$$

que en un diseño con sólo dos niveles por factor se reduce a:

$$AB|C_1 = \mu_{111} - \mu_{211} - (\mu_{121} - \mu_{221})$$

Utilizando el mismo argumento, la interacción entre los factores  $A$  y  $B$  en el nivel  $C_2$  viene dada por:

$$AB|C_2 = \mu_{jk2} - \mu_{j'k2} - (\mu_{jk'2} - \mu_{j'k'2})$$

que en un diseño con sólo dos niveles por factor se reduce a:

$$AB|C_2 = \mu_{112} - \mu_{212} - (\mu_{122} - \mu_{222})$$

Si no existe interacción de segundo orden, entonces los efectos simples de interacción  $AB|C_1$  y  $AB|C_2$  serán iguales. Por tanto:

$$\mu_{111} - \mu_{211} - (\mu_{121} - \mu_{221}) = \mu_{112} - \mu_{212} - (\mu_{122} - \mu_{222}) \quad [2.9]$$

que no es más que la expresión [2.7] referida a un diseño  $2 \times 2 \times 2$ . Así pues, definir la interacción a partir de las *desviaciones que experimentan las medias de las casillas respecto de sus correspondientes medias marginales* (primera definición; ecuación [2.6]) es equivalente a definirla a partir de las *diferencias entre las interacciones de primer orden en los diferentes niveles del tercer factor* (segunda definición; ecuación [2.7]). Y, de forma similar a como ocurría con las interacciones de primer orden, las de segundo orden son más fáciles de entender cuando se basan en [2.7].

En un diseño  $2 \times 2 \times 2$ , la interacción de segundo orden sólo tiene un grado de libertad. Por tanto, si el estadístico  $F$  asociado al efecto de la interacción es significativo, una interpretación basada en las diferencias incluidas en [2.9] agota por completo el significado de la interacción. Por tanto, para interpretar correctamente una interacción que tiene asociado un único grado de libertad, no es necesario recurrir a comparaciones adicionales para averiguar dónde se dan las diferencias: las diferencias se dan siguiendo la pauta propuesta en [2.9]. En este sentido, hay que tener presente que si el efecto de la interacción  $AB$  es el mismo en los dos niveles de  $C$ , entonces el efecto de la interacción  $AC$  es el mismo en los dos niveles de  $B$  y el efecto de la interacción  $BC$  es el mismo en

los dos niveles de  $A$ . Es decir, si se verifica la igualdad [2.9] referida a la interacción  $AB$ , también se verifica la igualdad:

$$\mu_{111} - \mu_{211} - (\mu_{112} - \mu_{212}) = \mu_{121} - \mu_{221} - (\mu_{122} - \mu_{222})$$

referida a la interacción  $AC$ , y la igualdad

$$\mu_{111} - \mu_{121} - (\mu_{112} - \mu_{122}) = \mu_{211} - \mu_{221} - (\mu_{212} - \mu_{222})$$

referida a la interacción  $BC$ .

En los diseños en los que alguno de los factores tiene más de dos niveles no se dan las mismas relaciones que en un diseño  $2 \times 2 \times 2$ . Si el efecto de la interacción tiene asociados dos o más grados de libertad, una  $F$  significativa indica que existe interacción pero no permite conocer su significado. En estos diseños, para poder interpretar correctamente el efecto de la interacción es necesario desarrollar comparaciones adicionales que, como se explica más adelante en el capítulo 5 “Cómo efectuar contrastes para interpretar la interacción”, permitan evaluar si los efectos simples de interacción entre dos factores son o no iguales en todos los niveles del tercer factor.

### 3. Errores cometidos al analizar y/o interpretar la interacción

Para cubrir el primer objetivo de este trabajo, se han revisado los 1.315 artículos publicados en *Psicothema* (861 artículos entre 2000 y 2006), el *International Journal of Clinical and Health Psychology* (190 artículos entre 2001 y 2006), *Psicologica* (121 artículos entre 2000 y 2006) y el *Spanish Journal of Psychology* (143 artículos entre 2000 y 2006). Se han elegido estas cuatro revistas por ser las revistas españolas de psicología con mayor impacto acumulado en los últimos años. *Psicothema* es una de las dos revistas españolas de psicología indexadas en el Journal Citation Reports; y el *International Journal of Clinical and Health Psychology* es la revista de psicología con mayor índice de impacto en los últimos años en la lista IN-RECS de revistas españolas de ciencias sociales (<http://ec3.ugr.es/in-recs/>). *Psicológica* y *The spanish Journal of Psychology* ocupan respectivamente las posiciones 7 y 12 en el índice de impacto acumulado 2002-2006, pero las revistas que ocupan las posiciones intermedias (3 a 7 y 8 a 11) son de corte clínico, por lo que no tiene interés incluirlas en esta revisión.

En una primera revisión se han seleccionado los artículos en cuyos resúmenes se indicaba que se había utilizado un ANOVA de dos o más factores y aquellos que en virtud del diseño informado requerían ese abordaje analítico. Puesto que los resúmenes no siempre facilitaban información suficiente para determinar si se había utilizado o no un ANOVA factorial o si el diseño exigía su utilización, se procedió a una revisión



visual del cuerpo de los artículos dudosos en busca de tablas o gráficos que sugiriesen la utilización de un diseño factorial.

Todos los artículos elegidos fueron impresos y leídos íntegramente para confirmar la utilización de un ANOVA factorial o la presencia de un diseño susceptible de ser analizado mediante la aplicación de un ANOVA factorial. En total se seleccionaron 272 artículos sobre los que se emitieron 365 juicios, ya sea porque en algún artículo se presentaba más de un estudio que merecía diferente clasificación, ya porque en un mismo experimento se presentaban resultados o afirmaciones que se clasificaban de diferente manera. Para 191 artículos se emitió un único juicio, para 70 artículos se emitieron 2 juicios, para 10 artículos se emitieron 3 juicios y para 1 artículo se emitieron 4 juicios. De los 365 juicios, 295 se refieren a interacciones entre 2 factores y 70 a interacciones entre 3 factores.

De los 272 artículos seleccionados, 167 pertenecen a la revista *Psicothema*, 25 al *International Journal of Clinical and Health Psychology*, 53 a *Psicologica* y 27 al *Spanish Journal of Psychology*. En el anexo 1 se presentan las citas bibliográficas de los 272 artículos seleccionados.

### **3.1. Errores cometidos al analizar y/o interpretar la interacción de primer orden**

Para estudiar los errores cometidos al analizar o interpretar la interacción vamos a comenzar con la interacción de primer orden o interacción entre dos factores. Esta interacción no sólo es más fácil de abordar que la de segundo orden (interacción entre tres factores), sino que se ha encontrado con mucha mayor frecuencia en los artículos revisados.

De los 272 artículos revisados, 269 contienen análisis de interacción de primer orden o diseños susceptibles de ser analizados de esta manera. Dependiendo de la forma de analizar e interpretar la interacción, los 295 juicios emitidos se han clasificado en una de las siguientes categorías:

- A) *No se analiza o no se interpreta la interacción*, a pesar de que los objetivos explícitos del estudio y/o el diseño utilizado lo requieren. Se incluyen aquí los artículos en los que únicamente se ofrecen resultados descriptivos (al margen del análisis de los efectos principales) y, consiguientemente, la interacción se interpreta a partir de un gráfico o de una tabla de medias (ocasionalmente se hacen afirmaciones en la discusión que sugieren la presencia de una interacción que no ha sido contrastada). También se incluyen aquellos estudios en los que existe una interacción significativa a la que no se presta atención. Y finalmente se incluyen aquellos estudios en los que la interacción se analiza de forma global, pero a la hora de interpretarla se recurre a una tabla de medias o gráfico cuando resultaría necesario recurrir a contrastes adicionales para hacer una interpretación adecuada.
- B) *La interacción se interpreta analizando los efectos simples*. Aquí se han incluido los artículos en los que, después de encontrar una interacción significativa, se recurre al análisis de los efectos simples para interpretarla, y aquellos en los que se recurre directamente al análisis de los efectos simples sin valorar previamente la presencia de una interacción significativa.

- C) *La interacción se analiza e interpreta correctamente*: ya sea comparando diferencias, ya sea argumentando que la combinación de efectos difiere de la suma de efectos.

Aunque varios de los artículos considerados incluyen diseños de más de 2 factores, en este apartado sólo se consideran los juicios que merecen las interacciones entre dos factores. La Tabla 3.1 ofrece el resultado de la clasificación. Puesto que algunos artículos incluyen más de un estudio, la suma de las clasificaciones llevadas a cabo es mayor (295) que el número de artículos revisados (269 pues en tres artículos no se hace alusión a interacción significativa de primer orden). En los artículos que contienen más de una interacción significativa, si todas ellas se interpretan de la misma manera se ha emitido un solo juicio; si hay distintas formas de interpretarla, se han emitido tantos juicios como formas de interpretarla. En los apartados que siguen se explica cada categoría de clasificación y las razones por las cuales las categorías A y B se consideran estrategias incorrectas o erróneas. Además, más adelante se presentarán ejemplos de cada tipo de error y de interpretaciones correctas (categoría C).

**Tabla 3.1. Clasificación de las interacciones entre dos factores en función del análisis e interpretación realizados por los investigadores**

Categoría (interpretación de la interacción entre 2 factores)	Juicios	%
A) No se analiza o no se interpreta la interacción	55	18,6
B) La interacción se interpreta recurriendo al análisis de los efectos simples	215	72,9
C) La interacción se analiza e interpreta correctamente	25	8,5
TOTAL	295	100,0

## **A) Error consistente en no analizar o no interpretar la interacción de primer orden**

Este error, consistente en no analizar la interacción o en efectuar un análisis incompleto, se puede dividir en tres subgrupos:

1. Los trabajos en que no se utiliza el estadístico  $F$  ni ninguna otra estrategia analítica para determinar si el efecto de la interacción es estadísticamente significativo. De todos los trabajos en los que se da esta omisión cabe distinguir entre los que simplemente no prestan atención a la interacción, es decir, los que ni la analizan ni la interpretan, y los que, aun sin analizarla, la interpretan a partir del gráfico de líneas o del tamaño de las medias de las casillas.
2. Los trabajos que, habiendo analizado la interacción y habiendo resultado significativa, no la interpretan, es decir, no le prestan atención.
3. Los trabajos en los que habiéndose realizado un análisis global de la interacción, a la hora de interpretarla se ha recurrido al gráfico de líneas o al tamaño de las medias de las casillas omitiendo la realización de contrastes adicionales (necesarios para interpretar correctamente la interacción por tratarse de diseños con más de dos niveles en alguno de los factores).

Con la única excepción de las escasas ocasiones en las que las hipótesis de un estudio se refieren a comparaciones planeadas, no prestar atención a la interacción en un diseño factorial (lo que se hace en los trabajos clasificados en la categoría A de la Tabla 3.1) constituye un desacierto evidente: dado que una interacción significativa está indicando

que el efecto de un factor no es el mismo en todos los niveles del otro, puede afirmarse que el significado de los efectos principales de un diseño queda matizado (incluso alterado) por la presencia de una interacción significativa.<sup>1</sup> La postura más generalizada recomienda no prestar atención a los efectos principales en presencia de una interacción significativa (Games, 1973, pág. 305). Kirk (1995, pág. 370) afirma:

“... a significant interaction is a red flag signalling that tests of treatment provide only a partial and often misleading picture of what has happened in an experiment”

y en el mismo sentido se expresan otros autores como Maxwell y Delaney (2004, pág. 301), Pedhazur y Pedhazur (1991, pág. 523), Winer y cols (1991, págs. 326-327), etc.

Por tanto, no prestar atención a la interacción en un diseño factorial (bien porque no se analiza, bien porque no se interpreta) supone cometer un error consistente en desechar información que no sólo es relevante en sí misma, sino que, además, podría estar modificando sensiblemente el significado de los efectos principales.

Por otro lado, interpretar la interacción a partir de las medias muestrales (o de la representación gráfica de esas medias) sin valorar su significación estadística supone olvidar un aspecto crucial del análisis de datos: los resultados muestrales podrían estar reflejando, no diferencias reales, sino simplemente fluctuaciones atribuibles al azar muestral (es decir, se podría estar considerando significativo un efecto que no lo es).

Cuando en un diseño factorial no se analiza una interacción se está perdiendo una parte de la información fundamental que se puede obtener a partir de dicho diseño. Otro tanto sucede cuando la interacción se analiza pero no se interpreta, pues sólo en los contados casos en que el interés recae en comparaciones planeadas la interacción tendrá

---

<sup>1</sup> En este sentido, algunos autores (Howell, 2002 [p. 432]; Keppel y Wickens, 2004 [p. 244]; León y Montero, 2001 [p. 165]) sugieren que, siendo significativa la interacción, todavía podría tener sentido interpretar los efectos principales en determinadas circunstancias. En concreto, cuando se da la *interacción ordinal*, interacción en la que las medias de un grupo están consistentemente por encima de las medias del otro, aunque no por ello dejan de señalar que la interacción debe ser interpretada.

un interés secundario. Además, una revisión de los objetivos explícitos de los estudios revela que no es infrecuente que los investigadores estén interesados en la interacción aunque no la analicen formalmente.

## **B) La interacción de primer orden se interpreta recurriendo al análisis de los efectos simples**

Este error es el que se repite con mayor frecuencia; consiste en interpretar la interacción recurriendo al análisis de los efectos simples. El 72,9 % de los juicios emitidos han sido clasificados en esta categoría.

Consideremos un diseño  $2 \times 2$ . En la ecuación [2.4] se ha definido la interacción como una diferencia de diferencias. Ahora bien, la diferencia  $\mu_{11} - \mu_{21}$  es el *efecto principal simple* (o *efecto simple*) del factor  $A$  en el nivel  $B_1$ ; y la diferencia  $\mu_{12} - \mu_{22}$  es el *efecto simple* de  $A$  en  $B_2$ . Por tanto, *afirmar que existe efecto de la interacción equivale a afirmar que el efecto simple de  $A$  en  $B_1$  difiere del efecto simple de  $A$  en  $B_2$* . Pero justamente a partir de aquí es donde surge el problema, porque la afirmación referida a la diferencia entre los dos efectos simples se interpreta con frecuencia de esta incorrecta manera: si al analizar los dos efectos simples del factor  $A$  se comprueba que uno de ellos es significativo y el otro no, se puede concluir que los efectos simples del factor  $A$  no son iguales en los dos niveles del factor  $B$  (lo cual, por otro lado, estaría justificando la presencia de una interacción significativa).

¿Por qué es incorrecta esta interpretación? Porque se está afirmando que son distintas dos cosas que no se han comparado; es decir, se está afirmando que el efecto simple de  $A$  en  $B_1$  difiere del efecto simple de  $A$  en  $B_2$  sin haber comparado formalmente entre sí ambos efectos simples. De hecho, uno de los dos efectos simples de  $A$  podría ser significativo y el otro no tanto si existe interacción significativa como si no; y ambos

efectos simples podrían ser significativos o no significativos tanto si existe interacción significativa como si no. Esta es una de las principales ideas sobre las que quiere llamar la atención este trabajo.

Para ilustrar estas afirmaciones a continuación se presentan dos ejemplos con datos preparados al efecto: en el primero, aunque el efecto de la interacción no es significativo, uno de los efectos simples es significativo y el otro no; mientras que en el segundo, el efecto de la interacción es significativo, pero ninguno de los efectos simples alcanza la significación estadística.

**Ejemplo 1.** Se trata de un diseño con 2 factores intersujetos,  $A$  y  $B$ , cada uno con dos niveles,  $A_1$  y  $A_2$  y  $B_1$  y  $B_2$ , respectivamente. En la Tabla 3.2 se presentan los datos brutos y en la Tabla 3.3 las medias de las casillas, las medias marginales y la media total.

**Tabla 3.2. Datos de un diseño  $2 \times 2$**

$A$	$B$	$Y$	$A$	$B$	$Y$
1	1	1,5	2	1	1,6
1	1	1,9	2	1	2,0
1	1	2,1	2	1	2,2
1	1	2,3	2	1	2,4
1	1	2,5	2	1	2,6
1	1	2,5	2	1	2,6
1	1	2,7	2	1	2,8
1	1	2,9	2	1	3,0
1	1	3,1	2	1	3,2
1	1	3,5	2	1	3,6
1	2	1,9	2	2	2,6
1	2	2,3	2	2	3,0
1	2	2,5	2	2	3,2
1	2	2,7	2	2	3,4
1	2	2,9	2	2	3,6
1	2	2,9	2	2	3,6
1	2	3,1	2	2	3,8
1	2	3,3	2	2	4,0
1	2	3,5	2	2	4,2
1	2	3,9	2	2	4,6

**Tabla 3.3. Medias correspondientes a los datos de la Tabla 3.2**

	$B_1$	$B_2$	Total
$A_1$	2,50	2,90	2,70
$A_2$	2,60	3,60	3,10
Total	2,55	3,25	2,90

$N = 10$  en todas las casillas; D.T. = 0,59 en todas las casillas

La tabla de ANOVA (Tabla 3.4) indica que los efectos principales son estadísticamente significativos, pero el efecto de la interacción,  $F_{(1, 36)} = 2,596$ , no alcanza la significación estadística ( $p = 0,116$ ).

**Tabla 3.4. Tabla de ANOVA**

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	$F$	Sig.
$A$	1,600	1	1,600	4,615	<b>0,038</b>
$B$	4,900	1	4,900	14,135	<b>0,001</b>
$A * B$	0,900	1	0,900	2,596	<b>0,116</b>
Error	12,480	36	0,347		
Total	19,880	39			

Sin embargo, el análisis de los efectos simples (Tabla 3.5) revela que un efecto simple es estadísticamente significativo y el otro no. Mientras que las diferencias entre  $B_1$  y  $B_2$  en  $A_1$  no son estadísticamente significativas, las diferencias entre  $B_1$  y  $B_2$  en  $A_2$  si lo son. En este escenario, una interpretación de la interacción basada en el análisis de los efectos simples nos llevaría a conclusiones equivocadas, incluso a la contradicción de afirmar la presencia de un efecto de la interacción que realmente no existe.

**Tabla 3.5. Análisis de los efectos simples**

Niveles del factor $A$	Diferencia $B_1 - B_2$	$p$
$A_1$	-0,400	<b>0,137</b>
$A_2$	-1,000	<b>0,001</b>



**Ejemplo2.** Se trata de un diseño con 2 factores intersujetos,  $A$  y  $B$ , el primero con dos niveles,  $A_1$  y  $A_2$ , y el segundo con tres niveles  $B_1$ ,  $B_2$  y  $B_3$ . En la Tabla 3.6 se presentan los datos brutos y en la Tabla 3.7 las medias de las casillas, las medias marginales y la media total.

**Tabla 3.6. Datos de un diseño  $2 \times 3$**

$A$	$B$	$X$	$A$	$B$	$X$	$A$	$B$	$X$
1	1	2,0	1	3	1,5	2	2	2,0
1	1	2,4	1	3	1,9	2	2	2,4
1	1	2,6	1	3	2,1	2	2	2,6
1	1	2,8	1	3	2,3	2	2	2,8
1	1	3,0	1	3	2,5	2	2	3,0
1	1	3,0	1	3	2,5	2	2	3,0
1	1	3,2	1	3	2,7	2	2	3,2
1	1	3,4	1	3	2,9	2	2	3,4
1	1	3,6	1	3	3,1	2	2	3,6
1	1	4,0	1	3	3,5	2	2	4,0
1	2	1,5	2	1	1,5	2	3	2,0
1	2	1,9	2	1	1,9	2	3	2,4
1	2	2,1	2	1	2,1	2	3	2,6
1	2	2,3	2	1	2,3	2	3	2,8
1	2	2,5	2	1	2,5	2	3	3,0
1	2	2,5	2	1	2,5	2	3	3,0
1	2	2,7	2	1	2,7	2	3	3,2
1	2	2,9	2	1	2,9	2	3	3,4
1	2	3,1	2	1	3,1	2	3	3,6
1	2	3,5	2	1	3,5	2	3	4,0

**Tabla 3.7. Medias correspondientes a los datos de la Tabla 3.6**

	$B_1$	$B_2$	$B_3$	Total
$A_1$	3,00	2,50	2,50	2,67
$A_2$	2,50	3,00	3,00	2,83
Total	2,75	2,75	2,75	2,75

$N = 10$  en todas las casillas; D.T. = 0,59 en todas las casillas

El análisis de varianza (Tabla 3.8) no encuentra efectos principales significativos de  $A$  o  $B$ , pero encuentra un efecto estadísticamente significativo de la interacción  $AB$ ,  $F_{(2, 54)} = 4,808$ ,  $p = 0,012$ . Sin embargo, el análisis de los efectos simples (Tabla 3.9) encuentra que ninguno de los tres efectos simples es estadísticamente significativo ( $p = 0,063$  para los

tres efectos de  $A$  en  $B_1$ ,  $B_2$  y  $B_3$ ), por lo que siguiendo la estrategia de explicar la interacción por medio del análisis de los efectos simples habría que concluir que la interacción encontrada en el contraste de interacción global no existe.

**Tabla 3.8. Tabla de ANOVA**

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	$F$	$p$
$A$	0,417	1	0,417	1,202	0,278
$B$	0,000	2	0,000	0,000	1,000
$A * B$	3,333	2	1,667	4,808	<b>0,012</b>
Error	18,720	54	0,347		
Total	22,470	59			

**Tabla 3.9. Análisis de los efectos simples**

Niveles del factor $B$	Diferencia $A_1 - A_2$	$p$
$B_1$	0,500	<b>0,063</b>
$B_2$	-0,500	<b>0,063</b>
$B_3$	-0,500	<b>0,063</b>

Los dos ejemplos presentados ilustran la inconsistencia existente entre el análisis de la interacción y el análisis de los efectos simples, pues la coexistencia de efectos simples de diferente interpretación (uno significativo y otro no) no garantizan una interacción significativa (ejemplo 1), mientras que la coexistencia de efectos simples todos ellos de similar significado (todos significativos o todos no significativos) es compatible con una interacción significativa (ejemplo 2).

La razón de la inconsistencia presentada en los ejemplos anteriores radica en el hecho de que un efecto simple incluye tanto el efecto principal como el de la interacción. Volviendo a la notación de la Tabla 2.1, cada efecto simple de  $A$  en  $B_1$  viene dado por (ver Kirk, 1995, págs. 377-378):

$$\alpha_j \text{ en } B_1 = \alpha_j + (\alpha\beta)_{j1} \quad [3.1]$$

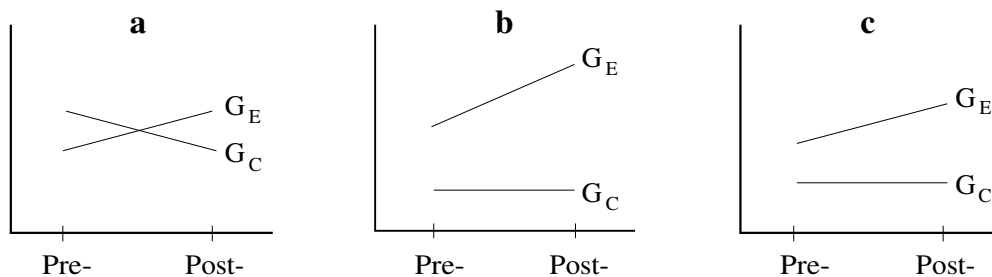
donde el término  $\alpha_j = \mu_{j+} - \mu$  se refiere a cada uno de los  $J$  efectos asociados a los niveles del factor  $A$  y el término  $(\alpha\beta)_{j1}$ , ya definido en la ecuación [2.1], se refiere a los  $J$  términos de interacción asociados al nivel 1 del factor  $B$ . De la ecuación [3.1] se deduce que un efecto simple puede ser significativo bien porque es significativo el efecto principal que incluye (es decir,  $\alpha_j \neq 0$  para algún  $j$ ), bien porque es significativo el efecto de la interacción (es decir,  $(\alpha\beta)_{j1} \neq 0$  para algún  $j$ ), bien por ambas cosas, bien porque  $\alpha_j + (\alpha\beta)_{j1} \neq 0$ . En palabras de Rosnow y Rosenthal (1989a), el efecto principal puede contribuir al efecto simple tanto o más que la interacción.

Las implicaciones prácticas de esta afirmación son importantes. Imaginemos que el factor  $A$  define dos grupos de tratamiento ( $G_E$  = experimental y  $G_C$  = control) y que el factor  $B$  representa dos momentos en el tiempo (pre- y postratamiento). En un diseño de estas características el investigador suele estar interesado en averiguar si el tratamiento tiene algún efecto sobre el grupo experimental (por supuesto, se asume que se trata de algún efecto distinto del no-tratamiento sobre el grupo control). Para obtener esta información no basta con analizar el efecto principal del factor  $A$ , sino que es necesario comparar lo que ocurre en el postratamiento (efecto simple de  $A$  en  $B_2$ ) con lo que ocurre en el pretratamiento (efecto simple de  $A$  en  $B_1$ ). Ahora bien, si para realizar esta comparación se recurre al análisis de los efectos simples por separado (estrategia utilizada en los trabajos clasificados en la categoría B de la Tabla 3.1), puede ocurrir que, *siendo significativo el efecto de la interacción*, no haya diferencias significativas entre  $G_E$  y  $G_C$  ni en el pre- ni en el postratamiento (ver Figura 3.1-a); y también puede ocurrir que haya diferencias significativas tanto en el pre- como en el postratamiento (ver Figura 3.1-b). En la estrategia basada en el análisis de los efectos simples por

separado, cualquiera de estos dos resultados llevaría a concluir que no es posible afirmar que exista efecto del tratamiento. Sin embargo, en clara discrepancia con esta conclusión, la presencia de una interacción significativa estaría indicando que la diferencia entre  $G_E$  y  $G_C$  no es la misma en el pre- y en el postratamiento; lo cual debería llevar a concluir que existe efecto del tratamiento (pues, en un diseño de estas características, una interacción significativa implica efecto del tratamiento).

También puede ocurrir que, *no siendo significativo el efecto de la interacción* (es decir, no habiendo diferencias entre lo que ocurre en el pre- y en el postratamiento), la diferencia entre  $G_E$  y  $G_C$  en el pretratamiento (efecto simple de  $A$  en  $B_1$ ) no sea significativa y sí lo sea la diferencia entre  $G_E$  y  $G_C$  en el postratamiento (efecto simple de  $A$  en  $B_2$ ). Este resultado podría llevar a afirmar que existe efecto del tratamiento cuando el hecho de que la interacción sea no significativa estaría descartando la presencia de un efecto del tratamiento (ver Figura 3.1-c).

**Figura 3.1. Distintas pautas de interacción en un diseño 2x2**



Por tanto, para poder afirmar que existe efecto del tratamiento no basta con saber que  $G_E$  y  $G_C$  no difieren en el pre- y sí en el postratamiento, como tampoco basta con saber que  $G_E$  cambia entre el pre- y el postratamiento mientras que  $G_C$  no lo hace (de todo esto es de lo que informan los efectos simples y donde con mayor frecuencia se centra el análisis de los investigadores). Para poder afirmar que existe efecto del tratamiento, la

diferencia observada en el post- hay que referirla a la observada en el pretratamiento (o, equivalentemente, el cambio observado en  $G_E$  entre el pre- y el postratamiento hay que referirlo al cambio observado en  $G_C$ ), y esto sólo es posible hacerlo comparando diferencias, que es justamente lo que se hace cuando se analiza el efecto de la interacción.

Aunque una interacción significativa coincidirá, en muchos casos, con la presencia de efectos simples diferenciados (es decir, unos significativos y otros no), esto no tiene por qué ser necesariamente así. Por tanto, si bien el análisis de los efectos simples podría llevar a las mismas conclusiones que la interpretación correcta de la interacción, esa estrategia debe ser considerada errónea porque puede conducir a conclusiones erróneas.

### **C) Interpretación correcta de la interacción de primer orden**

En este apartado se han clasificado los trabajos en los que: (1) la interacción se analiza correctamente comparando diferencias y (2) la interacción se interpreta correctamente ya sea comparando diferencias, ya sea argumentando que la combinación de efectos difiere de la suma de efectos.

La interpretación correcta de la interacción exige un abordaje diferente en diseños  $2 \times 2$  y en diseños en que alguno de los factores tiene más de dos niveles. En el primer caso basta con hacer una interpretación adecuada a partir de la tabla de medias o del gráfico de líneas, ya que la interacción implica las cuatro celdas disponibles, mientras que en el segundo es necesario realizar análisis adicionales para determinar cuáles son los niveles implicados en las diferencias significativas entre diferencias.

En la revisión bibliográfica realizada se han encontrado algunos trabajos con interpretaciones correctas de la interacción en diseños  $2 \times 2$ . Sin embargo, no se ha encontrado ninguna interpretación correcta asociada a diseños en los que alguno de los factores tenga más de dos niveles. Es decir, en ninguno de estos diseños se han realizado las comparaciones necesarias para analizar e interpretar correctamente una interacción significativa. Más adelante en el capítulo 5 “Cómo efectuar contrastes para interpretar la interacción” se explica como efectuar estas comparaciones.

### 3.2. Errores cometidos al analizar y/o interpretar la interacción de segundo orden

Los errores cometidos al analizar o interpretar la interacción entre tres factores se pueden clasificar en las mismas categorías ya utilizadas a propósito de la interacción entre dos factores, si bien la categoría referida a los efectos simples se debe subdividir en tres apartados atendiendo al tipo de análisis realizado para interpretar la interacción de segundo orden.

Dependiendo de la forma de analizar e interpretar la interacción, los 70 juicios emitidos se han clasificado en una de las siguientes tres categorías:

- A) *No se analiza o no se interpreta la interacción*, a pesar de que los objetivos explícitos del estudio y/o el diseño experimental lo requieren. Se incluyen aquí los estudios en los que únicamente se ofrecen resultados descriptivos, aquellos en los que existe una interacción significativa a la que no se presta atención y, finalmente, aquellos en los que la interacción se analiza de forma global, pero a la hora de interpretarla se recurre a una tabla o gráfico de medias cuando

resultaría necesario recurrir a contrastes adicionales para hacer una interpretación adecuada.

B) *La interacción se interpreta analizando los efectos simples.* Aquí se han incluido los estudios en los que una interacción significativa se intenta interpretar recurriendo al análisis de alguno de los diferentes tipos de efectos simples que se pueden considerar en una interacción entre tres factores:

B1) *Efectos simples simples:* analizan las diferencias entre los niveles de un factor en cada combinación de niveles de los otros dos factores.

B2) *Efectos simples:* analizan las diferencias entre los niveles de un factor en cada nivel del segundo factor tras agrupar los niveles del tercer factor.

B3) *Efectos simples de interacción:* analizan la(s) interacción(es) de primer orden dentro de cada nivel del tercer factor.

C) *La interacción se analiza e interpreta correctamente:* se comparan los efectos simples de interacción entre niveles distintos del tercer factor.

Aunque algunos de los artículos revisados incluyen diseños de más de 3 factores, en este apartado sólo se consideran los juicios que merecen las interacciones entre tres factores. La Tabla 3.10 ofrece el resultado de la clasificación de los 70 juicios emitidos. En los artículos que contienen más de una interacción significativa, si todas ellas se interpretan de la misma manera se ha emitido un solo juicio; si hay distintas formas de interpretarla, se han emitido tantos juicios como formas de interpretarla.

En los apartados que siguen se explica cada categoría de clasificación y las razones por las cuales las categorías A (no se analiza o no se interpreta la interacción) y

B (la interacción se interpreta recurriendo al análisis de los efectos simples) se consideran estrategias incorrectas o erróneas. Además, más adelante se presentan ejemplos de cada tipo de error y de interpretaciones correctas.

**Tabla 3.10 Clasificación de las interacciones entre tres factores en función del análisis e interpretación realizados por los investigadores**

Categoría (interpretación de la interacción entre 3 factores)	Juicios	%
A) No se analiza o no se interpreta la interacción	42	60,0
B) La interacción se interpreta recurriendo al análisis de los efectos simples	25	35,7
B1) Efectos simples simples	12	
B2) Efectos simples	7	
B3) Efectos simples de interacción	6	
C) La interacción se analiza e interpreta correctamente	3	4,3
TOTAL	70	100,0

### **A) Error consistente en no analizar o no interpretar la interacción de segundo orden**

Este error, consistente en no realizar un análisis de la interacción o en efectuar un análisis incompleto, se puede dividir en los mismos subgrupos considerados para la interacción entre dos factores:

- 1) Los estudios en que no se utiliza el estadístico  $F$  ni ninguna otra estrategia analítica para determinar si el efecto de la interacción es estadísticamente significativo.
- 2) Los estudios que, habiendo analizado la interacción y habiendo resultado significativa, no la interpretan, es decir, no le prestan atención. Sólo excepcio-



nalmente la omisión de la interpretación de una interacción significativa entre tres factores podría estar justificada cuando el interés del estudio recae en las posibles interacciones entre dos de los tres factores. Esta sería una situación de comparaciones planeadas “a priori”, donde el hallazgo de una interacción de segundo orden podría no tener interés.

- 3) Los estudios en los que habiéndose realizado un análisis global de interacción, a la hora de interpretarlo se ha recurrido al gráfico de líneas o al tamaño de las medias de las casillas omitiendo la realización de contrastes adicionales necesarios para poder interpretar correctamente la interacción. Esta interpretación sólo es correcta en diseños  $2 \times 2 \times 2$ ; en los diseños en que alguno de los factores tiene más de 2 niveles, la interpretación correcta exige recurrir a contrastes adicionales para determinar qué diferencias significativas entre interacciones de dos factores están involucradas en la interacción triple.

La falta de análisis o interpretación formal de la interacción entre tres factores es el hallazgo mas frecuente entre los juicios emitidos. Los comentarios aportados sobre el error consistente en no analizar o no interpretar la interacción entre dos factores son aplicables igualmente cuando dicho error se comete en relación con la interacción entre tres factores, por lo que no se comentan de nuevo.

### **B) La interacción de segundo orden se interpreta recurriendo al análisis de los efectos simples**

Al igual que ocurría en la interacción entre dos factores, el error consistente en interpretar la interacción recurriendo al análisis de los efectos simples también es frecuente cuando se trata de la interacción entre tres factores. Dicho error adopta diferentes

formas dependiendo del tipo de efectos simples que se considere para explicar la interacción. Los efectos simples de cualquiera de los tres tipos indicados, y que se discuten a continuación, pueden tener interés en sí mismos, pero no tienen capacidad para explicar una interacción significativa entre tres factores.

B1) Los *efectos simples simples* se refieren al efecto de un factor en cada una de las combinaciones de los otros dos factores. En la Tabla 3.11, la hipótesis nula referida a la interacción puede plantearse de la siguiente manera: la interacción entre  $A$  y  $B$  en  $C_1$  es igual a la interacción entre  $A$  y  $B$  en  $C_2$  e igual a la interacción entre  $A$  y  $B$  en  $C_3$ .<sup>2</sup>

Formalmente:

$$H_0: \mu_{111} - \mu_{211} - \mu_{121} + \mu_{221} = \mu_{112} - \mu_{212} - \mu_{122} + \mu_{222} = \mu_{113} - \mu_{213} - \mu_{123} + \mu_{223} \quad [3.2]$$

**Tabla 3.11. Notación en un diseño  $2 \times 2 \times 3$**

	$B_1$			$B_2$			
	$C_1$	$C_2$	$C_3$	$C_1$	$C_2$	$C_3$	
$A_1$	$\mu_{111}$	$\mu_{112}$	$\mu_{113}$	$\mu_{121}$	$\mu_{122}$	$\mu_{123}$	$\mu_{1++}$
$A_2$	$\mu_{211}$	$\mu_{212}$	$\mu_{213}$	$\mu_{221}$	$\mu_{222}$	$\mu_{223}$	$\mu_{2++}$
	$\mu_{+11}$	$\mu_{+12}$	$\mu_{+13}$	$\mu_{+21}$	$\mu_{+22}$	$\mu_{+23}$	$\mu$

Los *efectos simples simples* de  $A$  en niveles fijos de  $B$  y  $C$ , así como los *efectos simples simples* de  $B$  en niveles fijos de  $A$  y  $C$  comparan 2 celdas únicamente, mientras que los *efectos simples simples* de  $C$  en niveles fijos de  $A$  y  $B$  comparan los tres niveles de  $C$  en los niveles especificados de  $A$  y  $B$ . Estas comparaciones entre 2 ó 3 celdas no pueden dar cuenta de la interacción entre 2 ó más niveles de tres factores. Los *efectos simples simples* se definen de la siguiente manera:

<sup>2</sup> La hipótesis nula de interacción se puede plantear de otras formas, pero son todas equivalentes (ver apartado 2.2).

$$\begin{array}{lll}
\alpha_1 \text{ en } B_1C_1 = \mu_{111} - \mu_{+11} & \beta_1 \text{ en } A_1C_1 = \mu_{111} - \mu_{1+1} & \gamma_1 \text{ en } A_1B_1 = \mu_{111} - \mu_{11+} \\
\alpha_2 \text{ en } B_1C_1 = \mu_{211} - \mu_{+11} & \beta_2 \text{ en } A_1C_1 = \mu_{121} - \mu_{1+1} & \gamma_2 \text{ en } A_1B_1 = \mu_{112} - \mu_{11+} \\
& & \gamma_3 \text{ en } A_1B_1 = \mu_{113} - \mu_{11+} \\
\alpha_1 \text{ en } B_1C_2 = \mu_{112} - \mu_{+12} & \beta_1 \text{ en } A_1C_2 = \mu_{112} - \mu_{1+2} & \gamma_1 \text{ en } A_1B_2 = \mu_{121} - \mu_{12+} \\
\alpha_2 \text{ en } B_1C_2 = \mu_{212} - \mu_{+12} & \beta_2 \text{ en } A_1C_2 = \mu_{122} - \mu_{1+2} & \gamma_2 \text{ en } A_1B_2 = \mu_{122} - \mu_{12+} \\
& & \gamma_3 \text{ en } B_1C_2 = \mu_{123} - \mu_{12+} \\
\text{.....} & \text{.....} & \text{.....} \\
\alpha_j \text{ en } B_kC_l = \mu_{jkl} - \mu_{+kl} & \beta_k \text{ en } A_jC_l = \mu_{jkl} - \mu_{j+l} & \gamma_l \text{ en } A_jB_k = \mu_{jkl} - \mu_{jk+}
\end{array}$$

Y las hipótesis nulas correspondientes se pueden expresar como:

$$\begin{array}{ll}
H_0: \mu_{111} = \mu_{211} = \dots = \mu_{j11} = \mu_{+11} & \text{o bien: } \alpha_j \text{ en } B_1C_1 = 0 \text{ para todo } j \\
H_0: \mu_{112} = \mu_{212} = \dots = \mu_{j12} = \mu_{+12} & \text{o bien: } \alpha_j \text{ en } B_1C_2 = 0 \text{ para todo } j \\
H_0: \mu_{121} = \mu_{221} = \dots = \mu_{j21} = \mu_{+21} & \text{o bien: } \alpha_j \text{ en } B_2C_1 = 0 \text{ para todo } j \\
H_0: \mu_{122} = \mu_{222} = \dots = \mu_{j22} = \mu_{+22} & \text{o bien: } \alpha_j \text{ en } B_2C_2 = 0 \text{ para todo } j \\
\text{.....} & \text{.....} \\
H_0: \mu_{1kl} = \mu_{2kl} = \dots = \mu_{jkl} = \mu_{+kl} & \text{o bien: } \alpha_j \text{ en } B_kC_l = 0 \text{ para todo } j \\
\text{.....} & \text{.....} \\
H_0: \mu_{j1l} = \mu_{j2l} = \dots = \mu_{jkl} = \mu_{j+l} & \text{ó bien: } \beta_k \text{ en } A_jC_l = 0 \text{ para todo } k \\
\text{.....} & \text{.....} \\
H_0: \mu_{jkl} = \mu_{jk2} = \dots = \mu_{jkl} = \mu_{jk+} & \text{ó bien: } \gamma_l \text{ en } A_jB_k = 0 \text{ para todo } l
\end{array}$$

Es claro que ninguna de estas hipótesis capta la complejidad de la  $H_0$  expresada en [3.2]

B2) Los *efectos simples* consisten en agrupar los niveles de una de las variables independientes, dejando el diseño reducido a dos factores, y analizar los efectos simples de la misma manera que se hace en un diseño de dos factores. En la Tabla 3.12 se plantea un diseño similar al de la Tabla 3.11 pero, por conveniencia de la exposición, el factor que tiene tres niveles es el primero. Para analizar los *efectos simples* se agrupan

los niveles de la variable  $C$  presentando la notación correspondiente para, a partir de ella plantear las hipótesis nulas correspondientes a los *efectos simples*.

**Tabla 3.12. Notación en un diseño  $3 \times 2 \times 2$  en el que se ha agrupado el tercer factor ( $C$ ), celdas en fondo gris, para presentar la notación correspondiente**

	$B_1$			$B_2$			
	$C_1$	$C_2$		$C_1$	$C_2$		
$A_1$	$\mu_{111}$	$\mu_{112}$	$\mu_{11+}$	$\mu_{121}$	$\mu_{122}$	$\mu_{12+}$	$\mu_{1++}$
$A_2$	$\mu_{211}$	$\mu_{212}$	$\mu_{21+}$	$\mu_{221}$	$\mu_{222}$	$\mu_{22+}$	$\mu_{2++}$
$A_3$	$\mu_{311}$	$\mu_{312}$	$\mu_{31+}$	$\mu_{321}$	$\mu_{322}$	$\mu_{32+}$	$\mu_{3++}$
	$\mu_{+11}$	$\mu_{+12}$	$\mu_{+1+}$	$\mu_{+21}$	$\mu_{+22}$	$\mu_{+2+}$	$\mu$

Los *efectos simples* en el diseño agrupado son de la forma:

$$\begin{aligned}
 \alpha_1 \text{ en } B_1 &= \mu_{11+} - \mu_{+1+} & \beta_1 \text{ en } A_1 &= \mu_{11+} - \mu_{1++} \\
 \alpha_2 \text{ en } B_1 &= \mu_{21+} - \mu_{+1+} & \beta_2 \text{ en } A_1 &= \mu_{12+} - \mu_{1++} \\
 \alpha_3 \text{ en } B_1 &= \mu_{31+} - \mu_{+1+} & & \\
 \dots\dots\dots & & \dots\dots\dots & \\
 \alpha_j \text{ en } B_k &= \mu_{jk+} - \mu_{+k+} & \beta_k \text{ en } A_j &= \mu_{jk+} - \mu_{j++}
 \end{aligned}$$

Y las hipótesis nulas correspondientes se pueden expresar como:

$$\begin{aligned}
 H_0: \mu_{11+} = \mu_{21+} = \dots = \mu_{j1+} = \mu_{+1+} & \quad \text{o bien } \alpha_j \text{ en } B_1 = 0 \text{ para todo } j \\
 H_0: \mu_{12+} = \mu_{22+} = \dots = \mu_{j2+} = \mu_{+2+} & \quad \text{o bien } \alpha_j \text{ en } B_2 = 0 \text{ para todo } j \\
 \dots\dots\dots & \quad \dots\dots\dots \\
 H_0: \mu_{1k+} = \mu_{2k+} = \dots = \mu_{jk+} = \mu_{+k+} & \quad \text{o bien } \alpha_j \text{ en } B_k = 0 \text{ para todo } j \\
 H_0: \mu_{11+} = \mu_{12+} = \dots = \mu_{1k+} = \mu_{1++} & \quad \text{o bien } \beta_k \text{ en } A_1 = 0 \text{ para todo } k \\
 H_0: \mu_{21+} = \mu_{22+} = \dots = \mu_{2k+} = \mu_{2++} & \quad \text{ó bien } \beta_k \text{ en } A_2 = 0 \text{ para todo } k \\
 H_0: \mu_{31+} = \mu_{32+} = \dots = \mu_{3k+} = \mu_{3++} & \quad \text{ó bien } \beta_k \text{ en } A_3 = 0 \text{ para todo } k \\
 \dots\dots\dots & \quad \dots\dots\dots \\
 H_0: \mu_{j1+} = \mu_{j2+} = \dots = \mu_{jk+} = \mu_{j++} & \quad \text{ó bien } \beta_k \text{ en } A_j = 0 \text{ para todo } k
 \end{aligned}$$

Cada uno de los contrastes compara los niveles de uno de los factores en un nivel de otro factor cuando se agrupan los niveles del tercero, por lo que en ningún caso explican la interacción entre tres factores.

B3) Los *efectos simples de interacción* se refieren a la interacción entre dos factores en cada nivel del tercer factor. En este caso se trata de un auténtico contraste de interacción, pero que por sí mismo no puede explicar la interacción entre tres factores. En la Tabla 3.11 los *efectos simples de interacción*, expresados como diferencia entre casillas respecto a la diferencia entre los marginales correspondientes, son:

$$(\beta\gamma) \text{ en } A_1: (\mu_{111} - \mu_{121}) - (\mu_{11+} - \mu_{12+}) \quad \text{o bien: } (\mu_{112} - \mu_{122}) - (\mu_{11+} - \mu_{12+})$$

$$\text{o bien: } (\mu_{113} - \mu_{123}) - (\mu_{11+} - \mu_{12+})$$

$$(\beta\gamma) \text{ en } A_2: (\mu_{211} - \mu_{221}) - (\mu_{21+} - \mu_{22+}) \quad \text{o bien: } (\mu_{212} - \mu_{222}) - (\mu_{21+} - \mu_{22+})$$

$$\text{o bien: } (\mu_{213} - \mu_{223}) - (\mu_{21+} - \mu_{22+})$$

$$(\alpha\gamma) \text{ en } B_1: (\mu_{111} - \mu_{211}) - (\mu_{11+} - \mu_{21+}) \quad \text{o bien: } (\mu_{112} - \mu_{212}) - (\mu_{11+} - \mu_{21+})$$

$$\text{o bien: } (\mu_{113} - \mu_{213}) - (\mu_{11+} - \mu_{21+})$$

$$(\alpha\gamma) \text{ en } B_2: (\mu_{121} - \mu_{221}) - (\mu_{12+} - \mu_{22+}) \quad \text{o bien: } (\mu_{122} - \mu_{222}) - (\mu_{12+} - \mu_{22+})$$

$$\text{o bien: } (\mu_{123} - \mu_{223}) - (\mu_{12+} - \mu_{22+})$$

$$(\alpha\beta) \text{ en } C_1: (\mu_{111} - \mu_{121}) - (\mu_{+11} - \mu_{+21}) \quad \text{o bien: } (\mu_{211} - \mu_{221}) - (\mu_{+11} - \mu_{+21})$$

$$(\alpha\beta) \text{ en } C_2: (\mu_{112} - \mu_{122}) - (\mu_{+12} - \mu_{+22}) \quad \text{o bien: } (\mu_{212} - \mu_{222}) - (\mu_{+12} - \mu_{+22})$$

$$(\alpha\beta) \text{ en } C_3: (\mu_{113} - \mu_{123}) - (\mu_{+13} - \mu_{+23}) \quad \text{o bien: } (\mu_{212} - \mu_{223}) - (\mu_{+13} - \mu_{+23})$$

Y las hipótesis nulas correspondientes se pueden expresar como:

$$H_0: (\mu_{111} - \mu_{121}) = (\mu_{11+} - \mu_{12+}) \quad \text{o bien: } (\mu_{112} - \mu_{122}) = (\mu_{11+} - \mu_{12+})$$

$$\text{o bien: } (\mu_{113} - \mu_{123}) = (\mu_{11+} - \mu_{12+})$$

$$H_0: (\mu_{211} - \mu_{221}) = (\mu_{21+} - \mu_{22+}) \quad \text{o bien: } (\mu_{212} - \mu_{222}) = (\mu_{21+} - \mu_{22+})$$

$$\text{o bien: } (\mu_{213} - \mu_{223}) = (\mu_{21+} - \mu_{22+})$$

$$H_0: (\mu_{111} - \mu_{211}) = (\mu_{11+} - \mu_{21+}) \quad \text{o bien: } (\mu_{112} - \mu_{212}) = (\mu_{11+} - \mu_{21+})$$

$$\text{o bien: } (\mu_{113} - \mu_{213}) = (\mu_{11+} - \mu_{21+})$$

$$H_0: (\mu_{121} - \mu_{221}) = (\mu_{12+} - \mu_{22+}) \quad \text{o bien: } (\mu_{122} - \mu_{222}) = (\mu_{12+} - \mu_{22+})$$

$$\text{o bien: } (\mu_{123} - \mu_{223}) = (\mu_{12+} - \mu_{22+})$$

$$H_0: (\mu_{111} - \mu_{121}) = (\mu_{+11} - \mu_{+21}) \quad \text{o bien: } (\mu_{211} - \mu_{221}) = (\mu_{+11} - \mu_{+21})$$

$$H_0: (\mu_{112} - \mu_{122}) = (\mu_{+12} - \mu_{+22}) \quad \text{o bien: } (\mu_{212} - \mu_{222}) = (\mu_{+12} - \mu_{+22})$$

$$H_0: (\mu_{113} - \mu_{123}) = (\mu_{+13} - \mu_{+23}) \quad \text{o bien: } (\mu_{212} - \mu_{223}) = (\mu_{+13} - \mu_{+23})$$

Cada uno de los posibles contrastes que se pueden plantear a base de efectos simples de interacción compara los efectos simples de un factor entre los niveles del segundo factor en un nivel fijo del tercer factor, por lo que estos contrastes no pueden dar cuenta de la interacción entre los tres factores. Al igual que cualquier contraste de interacción entre 2 factores tiene que implicar como mínimo  $2 \text{ (niveles)}^2 \text{ (factores)} = 4$  celdas (Maxwell y Delaney, 2004), pues han de estar implicados al menos 2 niveles de cada uno de los factores, cualquier contraste de interacción que se refiera a tres factores tiene que implicar como mínimo  $2^3 = 8$  celdas.

### C) Interpretación correcta de la interacción de segundo orden

La interacción se analiza correctamente comparando las interacciones de primer orden entre los niveles del tercer factor. La interpretación correcta de la interacción tiene diferentes implicaciones en diseños  $2 \times 2 \times 2$  que en diseños en que alguno de los factores

tiene más de dos niveles. En el primer caso basta con hacer una interpretación basada en la tabla de medias o gráficos de líneas, ya que la interacción implica las ocho medias disponibles, mientras que en el segundo es necesario realizar análisis adicionales para determinar cuales son las medias implicadas en la interacción detectada. Mientras que en la revisión bibliográfica se han encontrado algunas interpretaciones adecuadas para diseños  $2 \times 2 \times 2$ , no se encontrado ningún artículo con diseños en que alguno de los factores tenga más de dos niveles en el que se haya realizado un desglose adecuado de la interacción entre tres factores. Más adelante se explicará como efectuar dichos contrastes.

### **3.3. Por qué se recurre al análisis de los efectos simples para interpretar una interacción significativa**

Como se ha señalado, la estrategia de recurrir al análisis de los efectos simples para interpretar una interacción significativa entre dos o tres factores está muy extendida. En nuestra opinión esto se debe al menos a dos razones. En primer lugar, a que las recomendaciones de los manuales de diseño y análisis no siempre son suficientemente claras y en segundo lugar a las limitaciones de los programas informáticos habitualmente utilizados por los investigadores en Psicología.

#### **3.3.1. Discrepancias entre manuales**

Diversos manuales de diseño y análisis estadístico han llamado la atención sobre el hecho de que los *efectos simples incluyen tanto efectos principales como de interacción*

(Kirk, 1995, págs. 377-378; Winer y cols, 1991, págs. 326-332), hasta el punto que Kirk (1995, pág. 383) ha llegado a afirmar:

“Test of hypotheses about simple main effects and contrasts that involve cell means, may be interesting, but they do not help us understand the interaction between two treatments”.

Esta afirmación ha sido recogida por otros autores como Jaccard (1998), pero a pesar de ello, prestigiosos manuales de diseño y análisis presentan los efectos simples como una estrategia apropiada (y en algunos casos única) para interpretar los datos en presencia de una interacción significativa (Howell, 2002, págs. 432, 489; Jaccard, 1998, pág. 20; Keppel y Wickens, 2004, pág. 247; Maxwell y Delaney, 2004, pág. 308; Pedhazur y Pedhazur, 1991, pág. 509; etc.). Algo parecido sucede también con las referencias en castellano (ver, por ejemplo, Ato y Vallejo, 2007, pág. 193, 198; León y Montero, 2001, pág. 165; Pascual, 1998, pág. 97).

A continuación se presenta una selección de comentarios y ejemplos que en nuestra opinión contribuyen a este estado de confusión.

Howell (2002, pág. 432) señala:

“...The analysis of simple effects can be an important technique for analyzing data that contain significant interactions. In a very real sense, it allows us to “tease apart” interactions.”

Como ilustración presenta un ejemplo, cuyas tablas y gráfico se reproducen a continuación. Se trata de un diseño 2 (*Age*: young; old)  $\times$  5 (*Condition*: counting, rhyming, adjective, imagery, intention). En la Tabla 3.13 se presentan los estadísticos descriptivos de la variable dependiente.



**Tabla 3.13. Estadísticos descriptivos (Tomada de Howell, 2002, pág. 432)**

	Condition					Mean
	Counting	Rhyming	Adjective	Imagery	Intention	
Old	7,0	6,9	11,0	13,4	12,0	10,06
Young	6,5	7,6	14,8	17,6	19,3	13,16
Mean	6,75	7,25	12,90	15,50	15,65	11,61

La tabla resumen del ANOVA (Tabla 3.14) muestra que los 2 efectos principales, *Age* y *Condition* son estadísticamente significativos, así como la interacción *Age*  $\times$  *Condition*. A continuación se presentan los efectos simples de *Condition* en cada nivel de *Age* y de *Age* en cada nivel de *Condition* que se reproducen en la Tabla 3.15.

**Tabla 3.14. Tabla de ANOVA (Tomada de Howell, 2002, pág. 432)**

Source	df	SS	MS	F
A (Age)	1	240,25	240,250	29,94*
C (condition)	4	1514,94	378,735	47,19*
A $\times$ C	4	190,30	47,575	5,93*
Error	90	722,30	8,026	
Total	99	2667,79		

\*  $p < 0,05$

**Tabla 3.15. Efectos simples (Tomada de Howell, 2002, pág. 432)**

Source	df	SS	MS	F
<b>Conditions (C)</b>				
C at Old	4	351,52	87,88	10,95*
C at Young	4	1353,72	338,43	42,15*
<b>Age (A)</b>				
A at Counting	1	1,25	1,25	<1
A at Rhyming	1	2,45	2,45	<1
A at Adjective	1	72,20	72,20	9,00*
A at Imagery	1	88,20	88,20	10,99*
A at Intention	1	266,45	266,45	33,20*
Error	90	722,30	8,03	

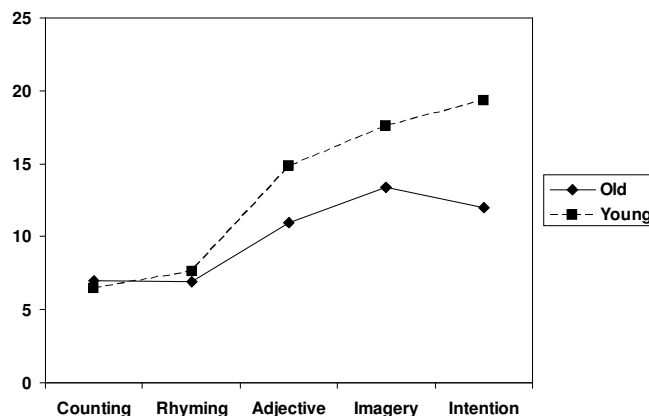
\*  $p < 0,05$

Howell concluye (pág. 435):

It is evident that differences due to conditions occur for both ages although the sum of squares for the older participants is only about one-quarter of what is for the younger ones. With regard to the Age effects, however, no differences occur in the lower-level tasks of counting and rhyming, but differences do occur on the higher-level tasks. In other words, differences between age groups show up on only those tasks involving higher level of processing. This is basically what Eysenck set out to demonstrate.

El análisis de los efectos simples indica dos cosas: 1) hay al menos dos tareas que difieren entre sí, tanto entre los jóvenes como entre los mayores, y 2) los jóvenes y los mayores difieren entre sí en tres de las cinco tareas (justamente en las que requieren un mayor nivel de procesamiento). Sin embargo, este tipo de análisis no aclara el significado de la interacción. Una interpretación correcta del efecto de la interacción exige comparar diferencias. La Figura 3.2 representa las medias de cada una de las celdas. Su inspección visual facilita una impresión que concuerda con el análisis de los efectos simples, pero tampoco representa una interpretación formal de la interacción. De hecho, si los efectos simples de *Age* en *Rhyming* y *Adjective*, *Imaginary* o *Intention* no fuesen significativamente diferentes, hipótesis que no se explora con el análisis de los efectos simples, el apoyo que estos resultados dan a la teoría de Eysenck se vería reducido.

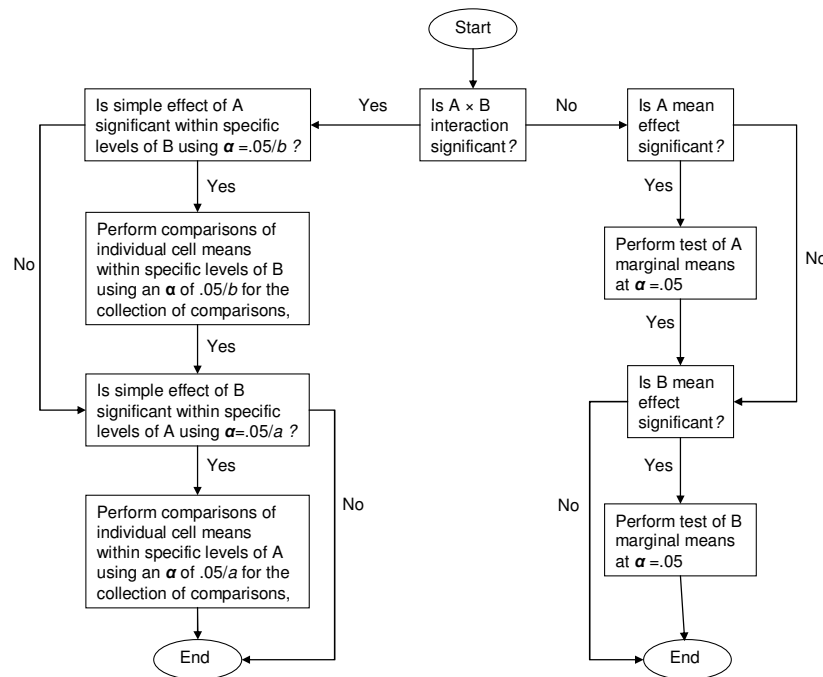
Figura 3.2. Tomada de Howell (2002, pág. 431)



Finalmente Howell comenta (pág. 435): “In general, we seldom look at simple effects unless a significant interaction is present”<sup>3</sup>

Maxwell y Delaney (2004), proponen a título orientativo el esquema general de trabajo presentado en la Figura 3.3.

**Figura 3.3 Tomada de Maxwell y Delaney (2004, pág. 308)**



General guideline for analyzing effects in a two-factor design.

En el esquema se hace hincapié en los efectos simples y en la comparación entre pares de casillas, que los autores explican de la siguiente manera:

When simple effects are performed in an attempt to interpret the meaning of a significant interaction, it is inevitable that multiple significance tests will be conducted...

<sup>3</sup> La excepción a la que se refiere Howell son los contrastes “a priori”, pero son muy pocos los estudios que hacen este tipo de planteamiento y además lo explicitan claramente en los objetivos del estudio.

...First consider tests for row effects –this is, tests of the factor A, which in our example is biofeedback. We will conduct three tests of biofeedback, one for each level of drugs [*drug* (B) has 3 levels]. The familywise  $\alpha$  can be maintained at .05 by performing each of these tests at an  $\alpha$  level of .05/3... Second consider test of the B factor.

In addition, if a simple effects test yields a significant result, it typically is necessary to test comparisons of individual cell means unless the factor in question has only two levels (in which case the precise nature of the differences is already identified).

En resumen, esta recomendación aconseja que una vez se haya encontrado una interacción significativa se pase a analizar los efectos simples, manteniendo el control sobre el nivel  $\alpha$  por familia de comparaciones, número de contrastes entre filas o número de contrastes entre columnas. Más adelante Maxwell y Delaney presentan un esquema similar, aunque de mayor complejidad para analizar la interacción entre tres factores (pág. 377). Estas recomendaciones tienen el problema de que si alguno de los factores tiene más de dos niveles el esquema propuesto no permite determinar con seguridad cuáles son las diferencias de medias implicadas en la interacción, pues el análisis de los efectos simples no informa sobre ello.

En 1973, Games sugirió que las preguntas teóricas que se hacen la mayoría de los investigadores son contestadas por los efectos simples. Esta afirmación resulta discutible, pues si tal fuese el caso no se desarrollarían estudios controlados, y hay que recordar que en el momento en que se diseña un estudio controlado la clave reside en el hallazgo de una interacción significativa, pues la pregunta que se desea responder es del tipo ¿lo que pasa en el grupo experimental es lo mismo que lo que sucede en el grupo control? Una cuestión diferente es que una vez establecido el efecto de la interacción, se desee estudiar *adicionalmente* la magnitud del efecto de la intervención sobre la

dependiente. El esquema general propuesto por Maxwell y Delaney resulta útil si una vez contrastada la interacción global el interés del investigador se centra en la magnitud de determinados efectos simples, pero no si su interés está en entender el significado de la interacción. Aunque estos autores previenen sobre la utilización estricta de la recomendación general y sugieren otros contrastes de interés entre los que incluyen aquellos que resultan útiles para desglosar la interacción en sus componentes, la aplicación de su esquema o estrategias similares se produce de forma acrítica en la mayoría de los estudios, como se pone de manifiesto en los artículos revisados en este trabajo. En contraposición a la afirmación de Games, Jaccard (1998, pág. 17), señala:

The vast majority of studies that use factorial designs in the social sciences pose research questions that focus on differences between mean differences, characterize interactions in terms of differences between mean differences, and draw substantive conclusions based on characterizations of differences between mean differences.

Un poco más adelante (pág. 20) advierte:

An analytical method frequently associated with interaction analysis is that of simple main-effects analysis. Numerous statisticians have noted that simple main-effects analysis has no necessary relationship with interaction analysis.

Sin embargo, más adelante (pág. 58) indica:

It is useful to supplement this [interaction] analysis with simple main-effects contrasts... However, a researcher often will want to know whether there is an effect of an independent variable on a dependent variable at each level of the moderator variable, and simple main-effects analysis provides perspectives on this.

Es decir, una vez contrastada la existencia de una interacción significativa, puede resultar interesante analizar los efectos simples como información complementaria, pero en absoluto idéntica o como forma de interpretar la interacción.

Algo parecido sucede en los trabajos publicados en español. Así León y Montero (2001, pág. 159) señalan:

“...cuando hay una interacción significativa, debemos reinterpretar los resultados de las variables independientes en términos de efectos simples.”

Y más adelante (pág. 161):

“Al no ser las pendientes de los efectos simples iguales... Debemos informar de los efectos simples por separado”

El punto sobre el que desean llamar la atención los autores es que en caso de que el efecto de la interacción sea significativo, la interpretación de los resultados en términos de efectos principales no tiene sentido. Sin embargo, esta advertencia se suele interpretar por los investigadores en términos de “el paso siguiente a encontrar una interacción significativa es analizar los efectos simples”, lo cual como se ha señalado no es una forma de interpretar la interacción, por lo que tendrá sentido o no según el objetivo del estudio. Y cuando el objetivo del estudio sea interpretar una interacción, el análisis de los efectos simples es innecesario.

En resumen, la mayoría de los manuales consideran el análisis de los efectos simples no como la forma de interpretar la interacción sino como una forma de obtener información adicional de interés para los investigadores. Sin embargo, a la hora de recomendar una estrategia para analizar la interacción frecuentemente se decantan por este procedimiento. Tal sucede en Maxwell y Delaney (2004), Howell (2002), hasta cierto punto en Jaccard (1998), posiblemente porque a menudo se basan en ejemplos

con diseños  $2 \times 2$  para estudiar la interacción entre dos factores, ó  $2 \times 2 \times 2$  para estudiar la interacción entre tres factores. Como en estos diseños el efecto de interacción global contrastado en la tabla de ANOVA no requiere de desgloses posteriores, pues necesariamente están involucradas todas las celdas del diseño, se pasa al análisis adicional de los efectos simples, dejando en un plano secundario los contrastes que realmente permitirían desglosar la interacción en diseños de mayor complejidad.

### 3.3.2. Limitaciones de los programas informáticos

El programa informático que se utiliza en la inmensa mayoría de los artículos revisados (el SPSS) no permite realizar un análisis *post hoc* de la interacción (al menos, no de forma directa): permite, mediante sintaxis, analizar los efectos simples y las comparaciones por pares asociadas a los efectos simples (ver Pardo y Ruiz, 2005, págs. 374-375, págs. 380-385); también permite comparar diferencias relacionadas con la interacción en los diseños de medidas repetidas (ver Field, 2005, págs. 471-478; esta estrategia no ha sido utilizada en ninguno de los artículos revisados); pero no permite efectuar los contrastes que aíslan el efecto de la interacción en los diseños factoriales con grupos al azar. Con otros programas disponibles en el mercado (SAS, Minitab, STATA...) ocurre algo parecido.

Una revisión de los manuales de análisis de datos con el programa SPSS pone de manifiesto que la estrategia analítica recomendada por la mayoría de ellos para interpretar una interacción significativa es, simplemente, la inspección del gráfico de líneas o el análisis de los efectos simples. En la Tabla 3.16 se ofrece un resumen de los procedimientos propuestos en estos manuales para la interpretación de una interacción significativa.

En el capítulo 5 “Cómo efectuar contrastes para interpretar correctamente la interacción” veremos cómo se pueden plantear los contrastes que permiten desglosar una interacción significativa en diseños de dos y tres factores.

**Tabla 3.16. Recomendaciones de los manuales de análisis de datos con SPSS para interpretar la interacción significativa**

Autor	Año	Recomendaciones para interpretar una interacción significativa		
		Ninguna	Ef. Simples	Gráfico
Álvarez, M.	1998	×		
Álvarez, R.	1994	×		
Balluerka, N., Vergara, M. I., Arnau y Gras, J.	2002		×	
Bryman, A.	2001			×
Camacho, J.	2004	Medias (celdas)	×	×
Coakes, S. J.	2005		×	×
Ferrán, M.	2001	×		
Gerber, S. B., Voelkl K.	2005	×		
Ho, R.	2006		×	
Kinnear, P. R., Gray C. D.	2006		×	
Landau, S., Everitt, B. S.	2004			×
Martín, P. J.	2006			
Martínez, M. A., Sánchez, A., Faulín, F. J. (eds.)	2007		×	×
Miller, R. L.; Campling, J. (ed.)	2002			×
Milton, J. S.	2007		×	
Pardo, A y Ruiz, M.A.	2005		×	×
Pérez, C.	2005			×
Pérez, C.	2005			×
Pérez, C.	2004			×
Rodríguez, M. J., Mora, R.	2001			×
Vicente, M. L.	2005		×	×
Visauta, B., Martorí i Cañas, J. C.	2003			×









## **4. Ejemplos de errores que se cometen al analizar la interacción**

En este capítulo se presentan ejemplos comentados del tipo de errores que se cometen al analizar y/o interpretar la interacción entre dos y tres factores. También se presentan algunos ejemplos en los que la interacción se analiza e interpreta correctamente

### **4.1. Ejemplos de errores que se cometen al analizar la interacción de primer orden**

Los ejemplos seleccionados se presentan agrupados en las tres categorías utilizadas en la Tabla 3.1: A) Error consistente en no analizar o no interpretar la interacción de primer orden; B) La interacción de primer orden se interpreta recurriendo al análisis de los efectos simples; y C) Interpretación correcta de la interacción de primer orden.

#### **A) Ejemplos de errores consistentes en no analizar o no interpretar la interacción de primer orden**

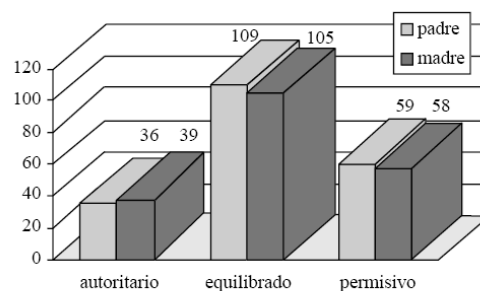
Se ha encontrado un único trabajo en que se concluye la existencia de la interacción entre dos factores sin realizar ningún análisis formal.

**Ejemplo 1) (Psicothema, 2005)** Los autores estudian en una escala de 0 a 150 puntos la autopercepción de los padres en las prácticas educativas, considerando el sexo de estos y tres posibles estilos educativos familiares. Se trata de un diseño con dos factores: uno intrasujetos, *estilo educativo familiar* (con tres niveles: autoritario, equilibrado, permisivo) y otro intersujetos, *sexo parental* (con dos niveles: padre, madre), siendo la variable dependiente la puntuación asignada a cada tipo de estilo educativo. Parte de los resultados se presentan en la Figura 1 del artículo, que se ha reproducido en la Figura 4.1. En el apartado de resultados (pág. 79) se describen los hallazgos como sigue:

*“Valoración de los padres: tanto en el padre como en la madre, la puntuación media más alta corresponde al estilo equilibrado, con gran diferencia con los otros dos estilos educativos, que resultaron distribuidos de tal manera que la media más baja es la que corresponde al estilo autoritario y la intermedia al permisivo (Figura 1). Sin embargo, hay una pequeña diferencia entre el padre y la madre, resultando las puntuaciones medias de la autopercepción del padre más elevadas en el estilo equilibrado y menos en el estilo autoritario” (cursiva mía).*

El párrafo anterior es una mera descripción de la figura en la que se presentan los resultados muestrales. A pesar de no mediar ningún análisis formal, se describe una interacción entre el sexo parental y los estilos educativos, centrándola en las diferencias entre sexo y los estilos equilibrado y autoritario.

**Figura 4.1. Tomada del original (Psicothema, 2005)**



*Figura 1. Percepción de padres y madres de lo que ellos creen que harían en cada estilo educativo familiar. Puntuación mínima 0 y máxima 150*

Este mismo error se está produciendo cuando se concluye la existencia de interacción sobre la tabla de medias o gráfico, aunque exista un análisis preliminar de los efectos principales. Se trata del mismo error porque el análisis de los efectos principales no es informativo de la posible interacción por lo que, en lo que se refiere a ésta, estamos en la misma situación de interpretar exclusivamente sobre un gráfico. Sin embargo, los autores se sienten más autorizados a concluir la existencia de interacción por el hallazgo de efectos principales significativos, por lo que este tipo de interpretación errónea es más frecuente como ilustran los siguientes ejemplos.

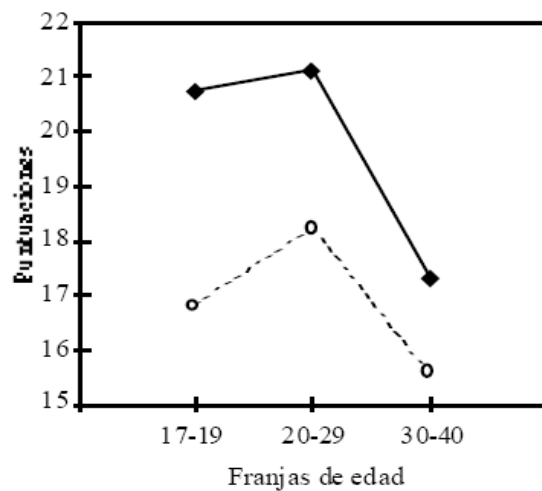
**Ejemplo 2)** (*Psicothema*, 2000) El autor estudia las diferencias que pueda haber en el rasgo de búsqueda de sensaciones en función de la edad y el sexo, utilizando la forma V de la Escala de Búsqueda de Sensaciones de Zuckerman. Se trata de un diseño con dos factores intersujetos: *grupo de edad* (con tres niveles: 17-19 a.; 20-29 a.; 30-40 a.) y *sexo* (con dos niveles: hombre, mujer). Respecto a las puntuaciones totales en la escala de Zuckerman, se presentan en dos tablas diferentes (Tablas 2 y 3 del artículo referenciado) las puntuaciones para ambos sexos ([media  $\pm$  D.T.]: hombres  $20,5 \pm 6,0$ ; mujeres:  $17,6 \pm 6,1$ ;  $t = 5,37$ ;  $p < 0,0005$ ) y para los tres grupos de edad considerados (17-19 años,  $18,3 \pm 6,1$ ; 20-29 años,  $19,6 \pm 6,2$ ; 30-40 años,  $16,6 \pm 6,5$ ;  $F = 7,87$ ;  $p < 0,0005$ ) En la Figura 4.2 se presentan los resultados desglosados por edad y sexo (Figura 2 del artículo referenciado) que el autor comenta de la siguiente manera (pág. 231).

*“Como se puede observar parece que la mayor diferencia en el rasgo entre hombres y mujeres, se presenta en la primera franja de 17-19 años y la menor diferencia en la franja de 30-40 años. Por otra parte, en la franja de 20-29 años se produce un aumento*

en la puntuación, siendo más notorio este aumento en la muestra de mujeres” (cursiva mía).

Los análisis realizados se refieren a los efectos principales de edad y sexo, pero en ningún caso a la interacción entre ambos factores, que se basa en la mera inspección visual del gráfico. Sin embargo, el comentario del autor se refiere al efecto de interacción.

**Figura 4.2** Tomada del original (*Psicothema*, 2000)



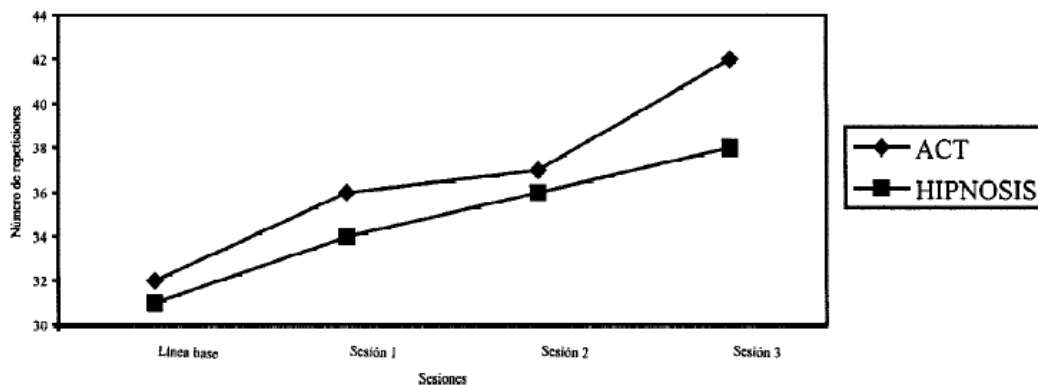
*Figura 2. Puntuaciones totales en función de la edad y el sexo (— hom - bres; ---- mujeres)*

**Ejemplo 3)** (*IJCHP*, 2004) Los autores estudian el efecto de la hipnosis y la terapia de aceptación y compromiso (ACT) en la mejora de la fuerza física de piragüistas. Se trata de un diseño con dos factores: uno intrasujetos, *sesiones* (con cuatro niveles: línea base; sesión 1; sesión 2; sesión 3), y otro intersujetos, *grupo de tratamiento* (con dos niveles: ACT, hipnosis). Los resultados respecto al número de repeticiones (fuerza muscular) se presentan en la la Figura 1 del artículo, que se reproduce en la Figura 4.3, y se presentan en el apartado resultados (pág. 489) de la siguiente manera:

“La variable intra-sujetos se refería a la paulatina aplicación de los tratamientos. Ambos grupos de sujetos obtuvieron mejoras estadísticamente significativas en el número de repeticiones al final de la intervención, en comparación con la línea base ( $F(1, 14) = 41,13; \alpha < 0,001$ ). *El rendimiento del grupo de ACT y el grupo de hipnosis corrieron parejos hasta llegar a la última medida, en que el grupo ACT mostró una ejecución visiblemente mejor que el grupo hipnosis.* Sin embargo, la variable entre-sujetos no llega a revelarse estadísticamente diferente a un nivel de confianza adecuado ( $F(1, 14) = 2,33; \alpha = 0,14$ )” (cursiva mía).

Los análisis realizados se refieren a efectos principales considerando sólo dos de los cuatro niveles intra-sujetos “...mejoras estadísticamente significativas en el número de repeticiones al final de la intervención, en comparación con la línea base...” y al efecto principal inter-sujetos. Ninguno de ellos informa sobre la interacción que de nuevo se basa en la mera inspección de la figura.

Figura 4.3. Tomada del original (IJCHP, 2004).



**Ejemplo 4) (Psicothema, 2002)** Los autores estudian la percepción de la calidad de vida en la infancia en función de la edad, el sexo y la nacionalidad de los niños encuestados. El diseño requiere un análisis con 3 factores inter-sujeto: el primero, *edad* (con dos niveles: 8 a.; 12 a.), el segundo, *sexo* (con dos niveles: hombre; mujer), y el tercero, *nacionalidad* (con dos niveles: española; argentina). En tanto que los investigadores tan



solo declaran el interés de estudiar la calidad de vida percibida por los niños en dos países diferentes y en dos grupos de edad diferentes, el análisis que realizan se circunscribe a dicha declaración de objetivos, comparando exclusivamente los dos niveles de cada uno de los factores por medio de la prueba T como se muestra en las Tablas 2, 3 y 4 del manuscrito referenciado (pág. 89) que se presentan agrupadas en la Tabla 4.1.

**Tabla 4.1. Tomada del original (*Psicothema*, 2002)**

<p><i>Tabla 2</i> Media y desviación típica de las puntuaciones en función del sexo</p>		
Factor	Media (DT)	
	Hombre	Mujer
Bienestar físico	3.95 (.63)	3.83 (.56)
Bienestar psíquico	3.81 (.53)	3.79 (.49)
Actividades vida diaria	3.58 (.55)	3.64 (.58)
Relaciones sociales	4.00 (.45)	4.00 (.49)

<p><i>Tabla 3</i> Media y desviación típica de las puntuaciones en función del país</p>		
Factor	Media (DT)	
	Argentina	España
Bienestar físico *	3.79 (.63)	4.02 (.54)
Bienestar psíquico *	3.69 (.56)	3.91 (.42)
Actividades vida diaria	3.55 (.58)	3.68 (.53)
Relaciones sociales	4.02 (.52)	3.98 (.39)
*Existen diferencias significativas entre los grupos, $p < 0.05$ .		

<p><i>Tabla 4</i> Media y desviación típica de las puntuaciones en función de la edad</p>		
Factor	Media (DT)	
	8 años	12 años
Bienestar físico*	3.76 (.53)	4.03 (.64)
Bienestar psíquico	3.74 (.59)	3.85 (.42)
Actividades vida diaria*	3.73 (.59)	3.48 (.51)
Relaciones sociales	4.04 (.45)	3.97 (.48)
*Existen diferencias significativas entre los grupos, $p < 0.05$ .		

Mientras que en el apartado de resultados se comentan cada una de las tablas de forma independiente del resto de los análisis, en la discusión (pág. 89), se dice:

*“Tanto en Argentina como en España las experiencias de satisfacción a la edad de 8 años se relacionan especialmente con...”* En el siguiente párrafo de la discusión leemos: *“A los 12 años la situación cambia notablemente... Los resultados difieren, por otra parte, entre Argentina y España”*.

Se está reivindicando una interacción a la que no se ha prestado atención ni en el diseño del estudio ni en el análisis de los resultados obtenidos y que por tanto no tiene ningún fundamento analítico.

**Ejemplo 5)** (*Psicothema*, 2001) Los autores desarrollan un estudio que según su propia declaración de objetivos *“...trata de explorar diferencias por sexo y edad en los estilos de personalidad del MIPS”* considerando dos grupos de edad: junior y senior. A pesar de ello, no plantean un diseño factorial con 2 factores intersujetos: el primero, *sexo* (con dos niveles: hombre; mujer), y el segundo, *edad* (con dos niveles: junior; senior), sino que estudian los dos factores de forma independiente. Los resultados se presentan en las Tablas 1 y 2 del artículo, que se reproducen en las Tablas 4.2 y 4.3, concluyéndose a partir de ellas (págs. 121):

*“De esta lectura inicial de las Tablas 1 y 2 se desprende que la edad es una fuente mayor de diferencias grupales que el sexo, ya que los valores del tamaño del efecto (d) aparecen más amplios y para más variables, en los sujetos clasificados como junior y senior que en los grupos de hombres y mujeres”*

Aunque en este párrafo no se reivindica la interacción entre edad y sexo, hay que hacerle dos críticas básicas: a) en ningún momento se han comparado directamente los efectos de la edad y el sexo sobre las dependientes; b) si ambos factores influyen en las dependientes parece obligado estudiar la posible interacción entre ellos.

Tabla 4.2. Tomada del original (*Psicothema*, 2001)

<p><i>Tabla 1</i> Diferencias entre grupos en cuanto al sexo. Medias, desviaciones típicas y valor de F en las diferencias entre grupos de sexo en las variables estudiadas, y valores del tamaño del efecto (d) y potencia observada en estos valores (P)</p>							
ESCALAS	MUJER N=289		VARÓN N=252		F	d	P
	Media	Des.T.	Varón	Des.T.			
BP (Bienestar Personal)	25,04	6,01	25,59	5,75	1,56		
AS (Adaptación Social)	25,13	5,21	24,69	5,48	0,79		
AP (Apertura)	21,79	7,99	22,83	7,78	2,28		
PR (Preservación)	19,86	10,31	17,69	10,16	6,81**	0,21	0,69
MO (Modificación)	33,3	4,28	33,83	4,47	1,96		
AC (Acomodación)	22,24	9,64	20,79	10,35	4,46*	0,14	0,39
SM (Si mismo)	17,28	8,09	21,11	8,37	29,39***	0,47	1
PO (Protección)	32,04	7,63	27,87	8,64	38,77***	0,51	1
EX (Extraversión)	26,85	9,25	24,56	9,1	6,39**	0,25	0,82
IT (Introversión)	10,8	7,52	12,7	7,39	8,42**	0,25	0,84
SE (Sensación)	17,17	6,83	17,76	6,59	0		
IN (Intuición)	23,86	8,97	22,83	8,64	0,18		
RF (Reflexión)	12,14	7,31	17,18	7,62	58,83***	0,68	1
AF (Afectividad)	32,69	7,89	27,71	8,77	49,08***	0,6	1
SI (Sistemización)	33	10,91	34,15	11,68	0,81		
IV (Innovación)	29,4	9,4	28,39	8,91	0,34		
RT (Retenimiento)	18,07	9,26	20,52	9,52	6,78**	0,26	0,86
CM (Comunicatividad)	31,15	11,12	31,92	11,22	1,45		
VA (Vigilación)	19,66	11,67	19,06	11,23	0,74		
FI (Firmeza)	30,3	10,33	33,53	10,43	14,25***	0,31	0,95
DI (Discrepancia)	23,35	9,26	24,46	7,96	2,74		
CF (Conformismo)	16,43	4,8	16,09	4,87	2,05		
SO (Sometimiento)	19,21	8,28	17,26	9	7,37**	0,23	0,74
CT (Control)	23,75	7,71	26,63	8,13	19,25***	0,36	0,99
IS (Insatisfacción)	26,71	9,43	26,27	9,15	0,41		
CC (Coordinación)	35,1	8,67	30,9	8,69	34,57***	0,48	1

\* p &lt; .05. \*\*p &lt; .01. \*\*\*p &lt; .001

Tabla 4.3. Tomada del original (*Psicothema*, 2001)

<p><i>Tabla 2</i> Diferencias entre grupos de edad. Medias, desviaciones típicas y valor de F en las diferencias entre grupos de edad en las variables estudiadas, y valores del tamaño del efecto (d) y potencia (P) en esos valores</p>							
ESCALAS	JUNIOR N= 325		SENIOR N= 216		F	d	P
	Media	Des.T.	Media	Des.T.			
BP (Bienestar Personal)	25,21	5,84	25,43	5,98	0,16		
AS (Adaptación Social)	25,07	5,32	24,71	5,36	0,48		
AP (Apertura)	21,99	8,05	22,71	7,68	0,9		
PR (Preservación)	18,3	10,75	19,68	9,51	2,87		
MO (Modificación)	33,51	4,59	33,61	4,04	0,02		
AC (Acomodación)	21,28	10,46	22	9,26	0,75		
SM (Si mismo)	18,96	8,73	19,21	7,98	0		
PO (Protección)	29,4	8,18	31,15	8,56	8,23**	0,21	0,66
EX (Extraversión)	27,3	9,11	23,51	8,99	20,85***	0,42	1
IT (Introversión)	10,95	7,44	12,79	7,49	7**	0,25	0,8
SE (Sensación)	15,31	6,46	20,64	5,78	93,58***	0,87	1
IN (Intuición)	25,88	8,41	19,62	8,08	71,68***	0,76	1
RF (Reflexión)	12,78	7,71	17,06	7,4	39,01***	0,57	1
AF (Afectividad)	30,93	8,62	29,52	8,7	2,27		
SI (Sistemización)	31,17	11,52	37,1	9,92	37,62***	0,55	1
IV (Innovación)	31,7	8,76	24,76	8,18	83,79***	0,82	1
RT (Retenimiento)	17,01	9,24	22,53	8,79	45,68***	0,61	1
CM (Comunicatividad)	33,15	11,33	29,03	10,43	18,67***	0,38	0,99
VA (Vigilación)	18,77	11,53	20,3	11,31	2,36		
FI (Firmeza)	31,69	10,53	31,98	10,46	0,01		
DI (Discrepancia)	24,49	8,96	22,94	8,19	4,56*	0,18	0,54
CF (Conformismo)	14,9	4,66	18,33	4,33	75,82***	0,76	1
SO (Sometimiento)	18,32	8,73	18,27	8,6	0		
CT (Control)	24,46	7,86	26,04	8,2	4,28*	0,61	0,2
IS (Insatisfacción)	26,26	9,69	26,88	8,68	0,63		
CC (Coordinación)	32,71	9,03	33,81	8,73	3,21		

\* p &lt; .05. \*\*p &lt; .01. \*\*\*p &lt; .001

Para concluir con los grupos de errores que tienen como factor común no analizar la interacción, es conveniente llamar la atención sobre una forma de eludir dicho análisis que consiste en combinar dos variables de clasificación en una sola *metavariante* con un número de niveles igual al producto del número de niveles en cada una de ellas. Esta forma de “*escamoteo de la interacción*” se completa con un ANOVA de un factor con una sola variable intersujetos (la *metavariante* creada artificialmente) y un análisis post-hoc de los pares de niveles que difieren entre sí.

**Ejemplo 6) (IJCHP, 2004)** Los autores estudiaron las diferencias psicosociales en personas portadores y no portadoras de VIH. Los grupos estudiados fueron portadores asintomáticos (hombres homosexuales, hombres heterosexuales y mujeres heterosexuales) y no portadores (hombres homosexuales, hombres heterosexuales y mujeres heterosexuales). Sin embargo, a la hora de realizar el análisis se crea “de facto” una meta-variable “*infección por VIH × sexo × condición sexual*”, con 6 niveles diferentes y se desarrolla un ANOVA de 1 solo factor intersujetos con 6 niveles. Los resultados se presentan en las Figuras 1 y 2 del artículo referenciado, que se reproducen en las Figuras 4.4 y 4.5. En el apartado resultados (pág. 62) se indica:

*“Se encontraron diferencias significativas en las variables depresión ( $F=4,733; p<0,002$ ) y ansiedad rasgo ( $F=4,863; p<0,001$ ) entre los distintos grupos, específicamente entre el grupo de hombres heterosexuales seropositivos en comparación con los tres grupos seronegativos, y en la variable ansiedad en los grupos de hombres y mujeres heterosexuales seropositivos en comparación con los grupos de hombres homosexuales y heterosexuales negativos (Figuras 1 y 2)”.*<sup>4</sup>

---

<sup>4</sup> Ni en el apartado Método ni en Resultados se indica como se realizaron las pruebas post-hoc

En la discusión se indica (pág. 63):

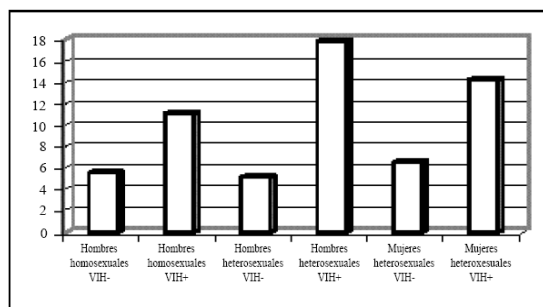
“Es necesario resaltar que esos síntomas (ansiedad) se presentan altos en ambos grupos de portadores, tanto en hombres como en mujeres heterosexuales. Parece relevante comentar que entre estos dos grupos no se encuentran diferencias, es decir, parece ser de acuerdo a nuestra muestra, *que el ser portador de VIH se relaciona con determinados grados de ansiedad independientemente de su condición sexual. Esto contrasta con la literatura disponible que considera que las mujeres sanas presentan mayores índices de ansiedad que los hombres sanos...*” (cursiva mía).

Posiblemente, un análisis de interacción *sexo × infección por HIV* habría resultado significativo, proporcionando evidencia para aclarar las dudas planteadas por el investigador en el párrafo anterior.<sup>5</sup> Finalmente, la discusión (pág. 64) acaba reivindicando una interacción entre sexo e infección por HIV sobre las alteraciones emocionales que realmente no se ha estudiado.

“Creemos que los datos aquí obtenidos demuestran que los sujetos que portan el virus presentan mayores índices de alteración emocional. De la misma forma, y adicional a lo anterior, *los síntomas se presentan en mayor medida en los sujetos heterosexuales y aún más en los hombres heterosexuales* (quienes presentan altos índices en las dos variables en las que se encontraron diferencias)” (cursiva mía).

**Figura 4.4. Tomada del original (IJCHP, 2004)**

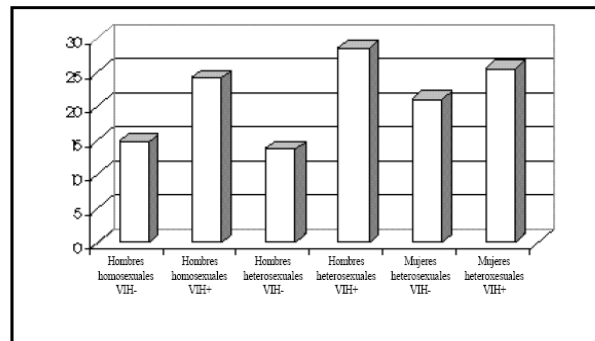
**FIGURA 1.** Distribución de medias de la variable depresión según grupos investigados.



<sup>5</sup> Hay que señalar que, como se verá más adelante, la realización de los contrastes adecuados por medio de la asignación de coeficientes a cada uno de los niveles de la metavariante permitiría realizar contrastes de interacción correctamente, aunque éste no es el caso del ejemplo.

**Figura 4.5. Tomada del original (IJCHP, 2004)**

**FIGURA 2.** Distribución de medias de la variable ansiedad-rasgo según grupos investigados.



**Ejemplo 7)** (*Psicothema*, 2002) Los autores estudiaron el efecto de la relajación muscular sobre la presión arterial y frecuencia cardiaca de pacientes hipertensos que recibían o no medicación para la hipertensión. Para ello distribuyeron al azar 40 pacientes entre cuatro grupos diferentes: relajación medicado, relajación no medicado, control medicado y control no medicado. La presión arterial y frecuencia cardiaca de todos los pacientes fue medida antes de la intervención, al término de la misma ocho semanas después, y en un seguimiento posterior 6 meses después. Por lo tanto se trata de un diseño factorial con tres factores: dos factores intersujetos, *técnicas de relajación* (con dos niveles: sí; no) y *medicación* (con dos niveles: sí; no), y un factor intrasujetos, *momento* (con tres niveles: pre-tratamiento; post-tratamiento; seguimiento). Parte de los resultados se presentan en la Tabla 3 del estudio, que se reproduce en la Tabla 4.4. En el apartado de resultados se indica (pág. 48):

“*Systolic blood pressure: Analysis of covariance revealed a treatment group difference on SBP ( $p < 0,000$ ) at post-treatment. Subsequent análisis [post hoc] using a Student «t» test revealed significant differences between the unmedicated relaxation group and unmedicated control group ( $p < 0,052$ ), between the medicated relaxation group and medicated control group ( $p < 0,006$ ) and between the unmedicated relaxation group*

*and medicated control group ( $p < 0,008$ )."* (cursiva mía). A Student "t" test showed that SBP decreased significantly in the unmedicated relaxation group ( $p < .000$ ) and medicated relaxation group ( $p < .01$ ). the unmedicated control group increased significantly SBP ( $p < .036$ )

En este ejemplo se desea centrar la atención en el efecto que tiene sobre el análisis la creación de la meta-variable, independientemente de otras consideraciones. Parece totalmente justificado estudiar la interacción *relajación*  $\times$  *medicación*, máxime cuando ambas intervenciones tuvieron un efecto sobre la presión arterial ambulatoria, pues es plausible pensar que el efecto de la relajación sobre la presión arterial será diferente en presencia o ausencia de medicación antihipertensiva. Sin embargo, como se ha creado una metavariante *relajación*  $\times$  *medicación* con cuatro niveles, no es necesario estudiar la posible interacción entre *relajación* y *medicación* sobre ninguna de las dependientes, SBP, DBP y HR; y bastaría con estudiar los pares de niveles que discrepan entre sí en la metavariante.

**Tabla 4.4 Tomada del original (*Psicothema*, 2002)**

Table 3 Ambulatory Blood Pressure Monitoring and Heart Rate Data (24 hours)			
Group	Pre-treatment mean (SD)	Post-treatment mean (SD)	Follow-up mean (SD)
(1) Unmedicated Relaxation			
SBP	141.5 (14.84)	127.7 (12.45)	128.2 (11.52)
DBP	90.1 (10.52)	80.6 (7.38)	78.8 (7.9)
HR	81.3 (12.95)	73.1 (5.74)	72.2 (6.76)
(2) Unmedicated Control			
SBP	134.3 (8.12)	138.2 (10.14)	141.2 (9.45)
DBP	86.5 (8.93)	87.5 (9.29)	89.5 (6.84)
HR	81.7 (8.12)	82.3 (7.12)	83.4 (6.6)
(3) Medicated Relaxation			
SBP	134.2 (12.73)	127.9 (11.23)	125 (9.78)
DBP	85 (10.31)	79 (7.62)	78 (7.62)
HR	76.8 (10.17)	71.9 (8.74)	71.5 (9.24)
(4) Medicated Control			
SBP	140.1 (16.99)	144.1 (12.03)	146.8 (11.02)
DBP	85.9 (10.18)	86.7 (11.15)	89.5 (11.08)
HR	71.6 (13.28)	74.4 (14.16)	77.1 (14.62)

Mientras que en los ejemplos presentados anteriormente no se ha estudiado la interacción formalmente, en los que se van a presentar a continuación sí que existe un análisis de interacción, si bien no se presta atención a los resultados o el análisis es incompleto y se interpreta informalmente sobre una figura. En primer lugar se presentarán ejemplos en los que se encuentra una interacción significativa que se pasa por alto y no se explica ni se tiene en cuenta a ningún efecto.

**Ejemplo 8)** (*Psicothema*, 2000) Los autores construyeron dos adaptaciones españolas del Test de Recuerdo Verbal Selectivo de las que estudiaron sus propiedades psicométricas. El apartado resultados (pág. 156), que se reproduce íntegramente a continuación, indica:

“Se realizó un ANOVA mixto con un factor intrasujetos (Forma) y dos factores intersujetos (sexo y orden de presentación) para cada variable dependiente, a fin de analizar los posibles efectos de estos factores, así como sus posibles interacciones. El nivel de significación nominal de 0,05 fue corregido utilizando el método de Bonferroni. El análisis de varianza indicó un efecto principal de la variable forma para RLTR ( $F(1, 46) = 13,98$ ;  $p < 0,002$ ) y una interacción significativa orden  $\times$  forma para Cued Recall [una de las variables dependientes] ( $F(1, 46) = 13,98$ ,  $p < 0,002$ )

En la Tabla 1 se presentan las medias, desviaciones estándar para cada forma, el coeficiente de correlación de Pearson y el coeficiente de correlación intraclase (ICC) de cada una de las medidas. En la Tabla 2 se presenta lo mismo para el orden 1 y en la Tabla 3 para el orden 2. intraclase” (tablas no reproducidas)

No se vuelve a nombrar dicha interacción en el apartado resultados, ni tampoco en la discusión, ni tan siquiera para explicar que el hallazgo no tiene interés.



**Ejemplo 9)** (*Psicothema*, 2005) Los autores estudiaron el efecto de presentar: a) estímulos intrusivos consistentes, asociados al mismo reforzador (grupo consistente); b) estímulos intrusivos inconsistentes, asociados a otro reforzador (grupo inconsistente); c) estímulos intrusivos asociados tanto al mismo como a otro reforzador (grupo no correlacionado); d) no presentar estímulos intrusivos (grupo control), con diferentes intervalos muestra-comparativo (EM-ECO's) de 0 a 16 s, siendo la variable dependiente el índice de discriminación. Se trata de un diseño con dos factores: uno intersujetos, *grupo* (con cuatro niveles: consistente; inconsistente; no-correlacionado; control), y otro intrasujetos, *intervalo EM-ECO's* (con cinco niveles: 0 s; 2 s; 4 s; 8 s; 16 s). En el primer párrafo de resultados (pág. 120) los autores indican:

*“La primera cuestión a determinar en el presente estudio fue si la intromisión de los estímulos intrusivos producía efectos distintos en la ejecución de los sujetos de cada uno de los grupos, específicamente sobre el índice de discriminación”* (cursiva mía)

Los resultados se presentan en la Figura 1 del artículo, que se reproduce en la Figura 4.6, y se comentan de la siguiente manera (pág. 120):

*“En la Figura 1 se presentan los índices de discriminación (ID's)...”* [sigue una descripción de la figura.] *“Un análisis de varianza de dos factores (grupo  $\times$  demora) mostró diferencias significativas entre grupos...”* [sigue una descripción del efecto principal de *grupo* con las comparaciones *post hoc* entre los cuatro niveles.] *“Adicionalmente se encontró una interacción estadísticamente significativa entre el valor del intervalo EM-ECO's y el grupo ( $F(12, 59) = 8,369, p < 0,001$ ).”* (cursiva mía)

En el resto del apartado de resultados no se vuelve a hacer mención de dicha interacción a pesar de haber resultado significativa.

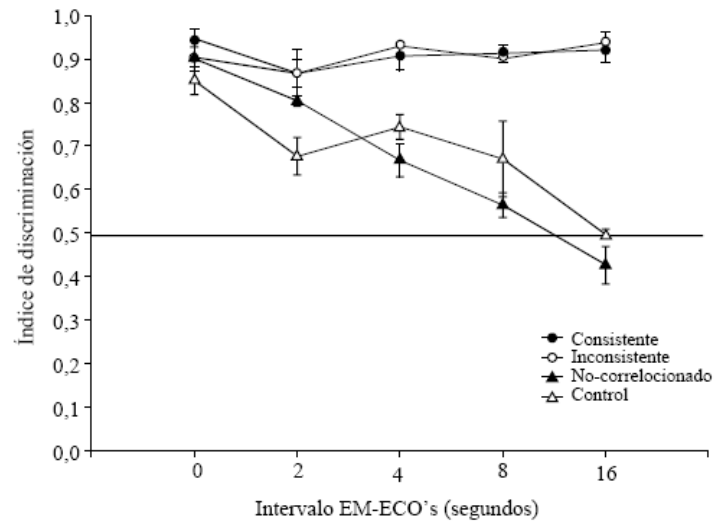
Figura 4.6. Tomada del original (*Psicothema*, 2005)


Figura 1. Índice de discriminación promedio de cada grupo en cada valor del intervalo EM-ECOS, la línea horizontal representa el nivel de azar

**Ejemplo 10)** (*Psicothema*, 2000) Los autores estudiaron la influencia de la ablación telencefálica y el reforzamiento (parcial o continuo) sobre el aprendizaje instrumental en la carpa dorada, tanto en la adquisición como en la extinción de la conducta de búsqueda de alimentación, en una de las cámaras de un laberinto acuático circular dividido en tres cámaras. Tanto para la adquisición como para la extinción se trata de un diseño de tres factores: un factor intrasujetos, *bloque de sesiones* (con 8 niveles), y dos factores intersujetos, *ablación telencefálica* (con dos niveles: ablación; cirugía de control), y *reforzamiento* (continuo [en todas las sesiones de adquisición]; parcial [en el 50% de las sesiones de adquisición]). Los resultados se presentan en la Figura 2 del artículo de referencia, que se reproduce en la Figura 4.7. En el apartado de resultados, después de hacer una breve descripción visual del periodo de adquisición se presentan el ANOVA de 3 factores, del que se indica (pág. 522):

“A Schedule (C, P)  $\times$  Lesión (L, S)  $\times$  4-Trial Block análisis of variance on the transformed latencies yielded a significant acquisition effect ( $F(7, 168) = 5,385, p <$

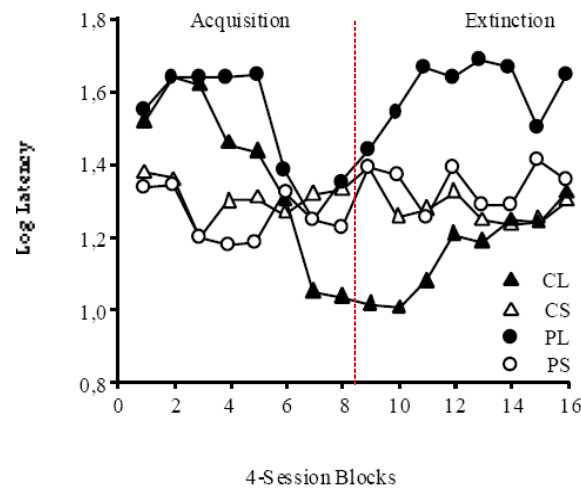
0,001), a significant *Lesion*  $\times$  *Trial* interaction ( $F(7, 168) = 5,835, p < 0,001$ ) and a significant main effect for lesion ( $F(1, 24) = 4,845, p < 0,04$ )” (cursiva mía).

En el apartado de resultados no se vuelve a citar esta interacción. Una situación similar se da en el periodo de extinción. De nuevo se comienza describiendo visualmente la parte de la figura correspondiente (niveles del 9 al 16) para continuar describiendo el ANOVA de tres factores, del que se dice:

“A Schedule (C, P)  $\times$  Lesión (L, S)  $\times$  4-Trial Block analysis of variance on the transformed latencies confirmed both observations. *There was a significant Lesion*  $\times$  *Trial* interaction ( $F(7, 168) = 2,292, p < 0,03$ ), and a significant main effect of schedule ( $F(1, 24) = 4,349, p < 0,05$ )” (cursiva mía).

En ambos casos el análisis de varianza se utiliza como un complemento de la inspección visual y no se le presta mayor interés.<sup>6</sup>

**Figura 4.7. Tomada del original (Psicothema, 2000)**



*Figure 2. Performance of groups of goldfish trained under either continuous (C) or partial (P) reinforcement during acquisition, and subject to telencephalic lesion (L) or sham operation (S) before the start of training.*

<sup>6</sup> NOTA relativa al *ejemplo 10*. Este ejemplo podría ser considerado igualmente como un caso en que después de encontrar una  $F$  de interacción significativa se acude a una tabla o figura para explicar la interacción (apartado siguiente). Sin embargo, la forma de presentar los resultados sugiere que primero se ha descrito la interacción sobre la figura y luego se ha utilizado la  $F$  de interacción exclusivamente para corroborar la impresión gráfica obtenida previamente.

A continuación se presentan ejemplos de estudios en los que la interacción no se analiza completamente: se ofrece un análisis formal de la interacción global y una interpretación informal. Se trata de estudios en los que, después de un contraste global de la interacción estadísticamente significativo, en lugar de desglosar e interpretar la interacción correctamente se trata de explicar su significado interpretando directamente una tabla o gráfico de medias. Este procedimiento, que sería correcto si cada uno de los factores no tuviese más que dos niveles, resulta inadecuado cuando alguno de los factores tiene tres o más niveles, puesto que la interpretación sobre la tabla o gráfico no garantiza entre qué niveles de los factores se están dando las diferencias que sustentan la interacción y a veces ni siquiera da pistas al respecto. El sustrato sobre el que se apoya este error consiste, al igual que alguno comentado anteriormente, en confundir los resultados obtenidos en la muestra (medias observadas) con las medias poblacionales, adjudicando a la población diferencias que pueden ser mera consecuencia del muestreo.

**Ejemplo 11** (*Psicothema*, 2000) Los autores estudiaron la influencia de la presión temporal y el canal de comunicación utilizado para el trabajo en grupo sobre la calidad del desempeño en tres tipos de tareas. Se trata de un diseño de dos factores intersujetos: el primero, *presión temporal* (con dos niveles: con presión; sin presión), y el segundo *canal de comunicación* (con tres niveles: cara a cara [CAC]; videoconferencia [VC]; correo electrónico [CE]). Los tres tipos de tarea fueron: intelectivas, de negociación y creativas. La variable dependiente, *calidad del desempeño*, se midió de diferente manera en función del tipo de tarea. En el apartado resultados (pág. 243) se señala:

“Se hallan diferencias significativas en función de la interacción *canal\*presión* en las tareas intelectivas ( $F = 3,98$ ,  $p = 0,021$ ) y en las creativas ( $F = 7,45$ ,  $p = 0,001$ ), aunque

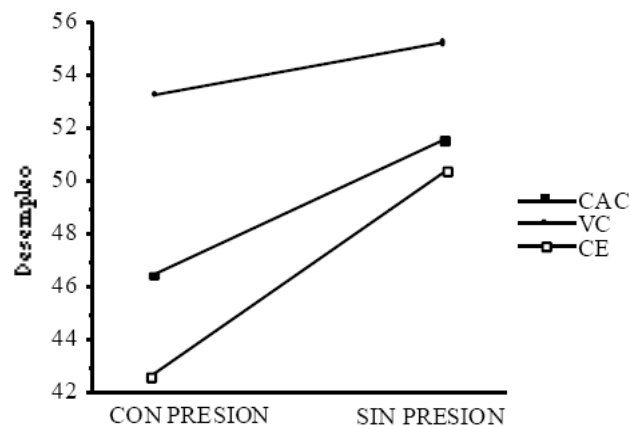
no en las de negociación ( $F = 0,54$ ,  $p = 0,584$ ). La representación gráfica de las medias nos ayuda a comprender los resultados.”

Los gráficos 1 y 3, que se reproducen en las Figuras 4.8 y 4.9, corresponden a las interacciones estadísticamente significativas. Respecto al gráfico 1 (véase la Figura 4.8), los autores señalan (pág. 243):

“Respecto a las tareas intelectivas, se observa que aunque el desempeño mejora en todos los canales de comunicación cuando trabajan sin presión temporal, *esta mejora es mucho mayor en el caso del correo electrónico y de cara a cara que en videoconferencia (ver gráfico 1). Dicho de otra manera, en las tareas intelectivas, la presión temporal afecta de manera más negativa a las condiciones de correo electrónico y de comunicación cara a cara que a la videoconferencia.* Por otra parte, como puede contemplarse en el gráfico 1, la presión temporal acentúa las diferencia en desempeño grupal en función del canal.” (cursiva mía)

Se está afirmando que existen diferencias que no se han constatado mediante el correspondiente análisis. Para confirmar, por ejemplo, que la mejora es “mucho mayor” en el caso del correo electrónico que en el de la videoconferencia es necesario comparar ambas mejoras, cosa que no se hace.

**Figura 4.8. Tomada del original (Psicothema, 2000)**



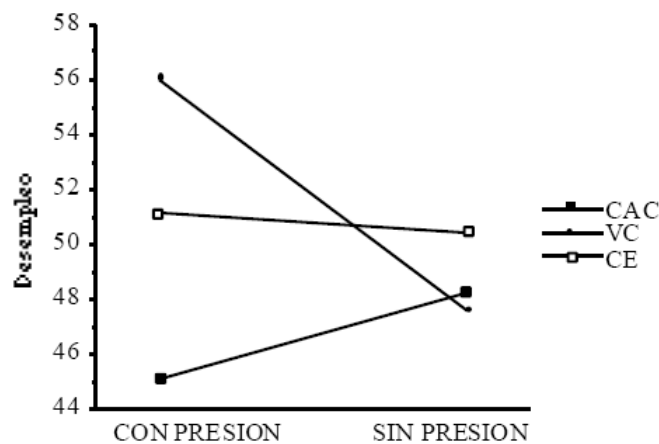
**Gráfico 1. Desempeño en tareas intelectivas en función de la interacción canal \* presión**

Respecto a su gráfico 3 (véase la Figura 4.9) los autores señalan (pág. 244):

*“Finalmente, en cuanto a las tareas creativas, la significación de las diferencias consiste en que mientras en correo electrónico y sobre todo en videoconferencia el desempeño es mejor en situaciones de presión temporal, en cara a cara el rendimiento es más alto cuando no existe presión temporal para la realización de la tarea.” (cursiva mía)*

De hecho en este caso no se llega a especificar entre qué niveles de canal se producen las diferencias entre efectos simples que justifican la interacción presión  $\times$  canal. Se da por supuesto que entre VC y CAC, pero no queda claro si también entre CE y CAC. De nuevo sería necesario desglosar formalmente la interacción por medio de contrastes lineales adecuados para averiguar entre qué niveles de canal se produce un efecto diferencial de la presión sobre el rendimiento medio.

**Figura 4.9. Tomada del original (*Psicothema*, 2000)**



**Gráfico 3. Desempeño en tareas creativas en función de la interacción canal \* presión**

**Ejemplo 12) (*Psicothema*, 2004)** Los autores estudiaron la relación entre el efecto de compatibilidad de los distractores y el efecto Simon sobre el tiempo de respuesta (TR) en el procesamiento de información. Además se utilizaron diferentes distancias objeto-

distractor (O-D). Todos los sujetos pasaron por todas las condiciones experimentales. Se trata de un diseño factorial con tres factores intrasujetos: *compatibilidad de los distractores* (con tres niveles: compatible; incompatible; neutro), *efecto Simon* [presentación de un estímulo acústico durante la emisión de la respuesta] (con tres niveles: mismo lado; lado contrario; ambos lados [la respuesta se tenía que emitir con la mano derecha o izquierda dependiendo de la información presentada en pantalla]) y *distancia O-D* (con dos niveles: 0,65°; 1,30°). En el apartado de resultados se indica:

“...resultaron estadísticamente significativas las siguientes [interacciones]: *estímulo acústico* × *compatibilidad de los distractores* ( $F(4, 44) = 5,074, p < 0,01, \eta^2 = 0,316$ )...”.

La Figura 4.10 reproduce la Figura 2 del artículo, donde se presentan los resultados, que se comentan en la pág. 278 como sigue:

“...en la Figura 2 se observa que mientras que en la condición de distractores compatibles apenas existe un efecto diferencial sobre los TRs según el tipo de estímulo acústico, en las condiciones incompatible y neutra aparece un fuerte efecto de interferencia del estímulo acústico, efecto que como veremos a continuación dependerá de la distancia que separa el objetivo de los distractores” (cursiva mía)

En este párrafo se interpreta la interacción por medio de la descripción de los efectos simples. Tal tipo de interpretación solo es posible en un diseño 2×2, donde la interacción captada por la *F* global involucra diferencias significativas entre los dos efectos simples definidos por las 4 celdas del diseño. Sin embargo, en un diseño con dos factores y tres niveles en cada factor como el del ejemplo, la interacción global se puede desglosar en 9 comparaciones, involucrando cada una de ellas las diferencias entre dos niveles de cada uno de los factores, de las que no se ha llevado a cabo ninguna.

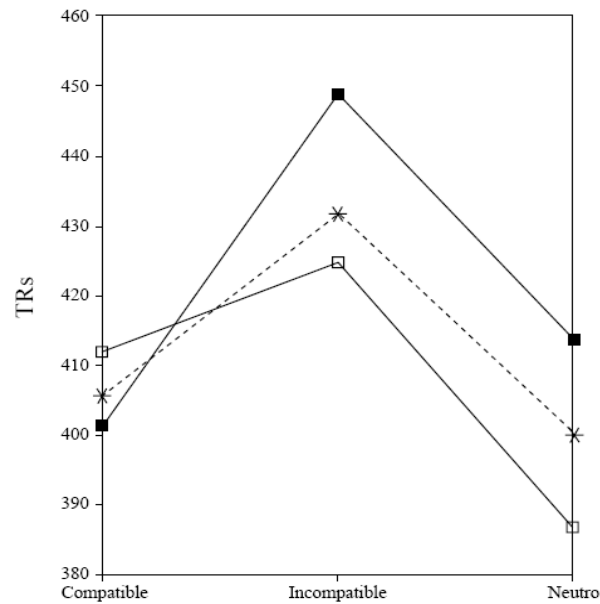
Figura 4.10. Tomada del original (*Psicothema*, 2004)


Figura 2. TRs en función de la compatibilidad de los distractores (compatible, incompatible, neutro) y del estímulo acústico: (□) mismo lado, (■) lado opuesto y (\*) ambos lados

## B) Ejemplos de errores en los que la interacción de primer orden se interpreta recurriendo al análisis de los efectos simples

Un tipo de error frecuente se produce cuando se analizan los efectos simples como medio para estudiar la interacción sin analizar ésta por medio de los contrastes adecuados. Este tipo de error tiene su fundamento en una interpretación equivocada de la interacción consistente en creer que hay interacción cuando uno o más efectos simples son significativos y uno o más efectos simples no los son. Consecuentemente con esta forma de pensar, es frecuente que los investigadores acudan directamente a los efectos simples sin realizar un verdadero análisis de interacción, que sería considerado como un “paso intermedio” de escaso interés. Este error se produce incluso cuando el



objetivo del estudio es poner de manifiesto una interacción entre dos factores o cuando se concluye la existencia de una interacción que no se analizó.

**Ejemplo 13)** (*Psicothema*, 2001) Los autores estudiaron el efecto de la asimetría hemisférica, presentación en el campo visual derecho o izquierdo, en el reconocimiento de palabras (se presentan tanto palabras como pseudopalabras), analizando la influencia de la frecuencia de utilización e imaginabilidad de las palabras. Si se consideran tan solo los aciertos, reconocimiento de palabras como tales, nos encontramos ante un diseño de tres factores intrasujetos: *campo visual* (con dos niveles: CVD-HI; CVI-HD), *frecuencia de las palabras* (con dos niveles: frecuentes; infrecuentes) e *imaginabilidad de las palabras* (con dos niveles: alta imaginabilidad [palabras concretas]; baja imaginabilidad [palabras abstractas]).

En lo referente al análisis del TR de las palabras, después de la presentación de los efectos principales de los factores *frecuencia de las palabras*, *imaginabilidad de las palabras* y *campo visual*, se dice (pág. 26):

*“La interacción entre el factor Frecuencia y el factor Imaginabilidad resultó estadísticamente significativa, ya que en las palabras frecuentes, los tiempos de las concretas son similares a los de las abstractas; sin embargo, en las palabras infrecuentes, las palabras abstractas tienen tiempos mucho mayores que las concretas [ $F(1, 36) = 32,91, p = ,000$ ]”* (cursiva mía)

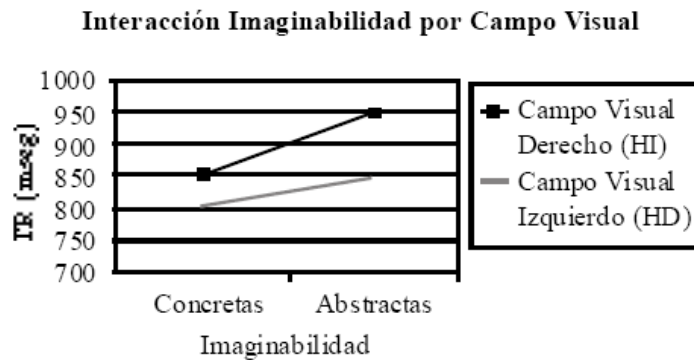
Este párrafo ilustra perfectamente el punto de vista según la cual la interacción es la consecuencia del hallazgo de efectos simples con diferente significación.

La Figura 4.11 se corresponde con la Figura 1 del artículo de referencia y es utilizada por los autores para describir la interacción *Imaginabilidad*  $\times$  *Campo Visual* de la siguiente manera (pág. 26):

*“También fue significativa la interacción imaginabilidad por campo visual (como puede observarse en la Figura 1) debido a que la diferencia entre CVD y CVI aumenta significativamente cuando se trata de palabras abstractas. Así, el TR que requiere el CVD-HI para las palabras abstractas es de 946 mseg y, en cambio, de 846 mseg en el CVI-HD [F (1, 36)=5,49, p = 0,025]” (cursiva mía)*

Lo que se está diciendo en este párrafo es que uno de los efectos simples no alcanza la significación estadística mientras el otro sí, razón por la cual se considera que existe interacción. La ausencia del análisis de interacción propiamente dicho impide saber si realmente dicha interacción existe o no (aunque la observación del gráfico inclina a pensar que no).

**Figura 4.11 Tomada del original (*Psicothema*, 2001)**



*Figura 1. Tiempo medio de Reacción (en mseg) de la interacción Imaginabilidad (palabras Concretas/Abstractas) por Campo Visual en el que es presentado el estímulo (Campo Visual Derecho (Hemisferio Izquierdo)/Campo Visual Izquierdo (Hemisferio Derecho)).*

**Ejemplo 14)** (*Psicothema*, 2003) El autor estudió la influencia de un programa de instrucción en la escritura en procesos cognitivos en niños con dificultades de aprendizaje de la escritura (DAE) que fueron asignados al grupo experimental o al grupo control y se introdujo un grupo control positivo consistente en alumnos sin DAE. Los niños fueron evaluados antes de aplicar el programa y después de la intervención (cinco meses después), realizándose un seguimiento al inicio del siguiente curso

académico (tres meses después). Se trata de un diseño de dos factores: con un factor intersujetos, *grupo* (con tres niveles: experimental; control negativo [sin intervención]; control positivo [con intervención]) y un factor intrasujeto, *momento* (con tres niveles: pretest; posttest; retest). En la Tabla 4 del apartado de resultados, que se reproduce en la Tabla 4.5, se presentan los análisis *post-hoc* con la prueba de Scheffé (pretest; posttest; retest) en cada uno de los niveles de la variable *grupo*, es decir, analizando los efectos simples de la variable intrasujeto en cada nivel de la variable intersujeto.

**Tabla 4.5. Tomada del original (*Psicothema*, 2003)**

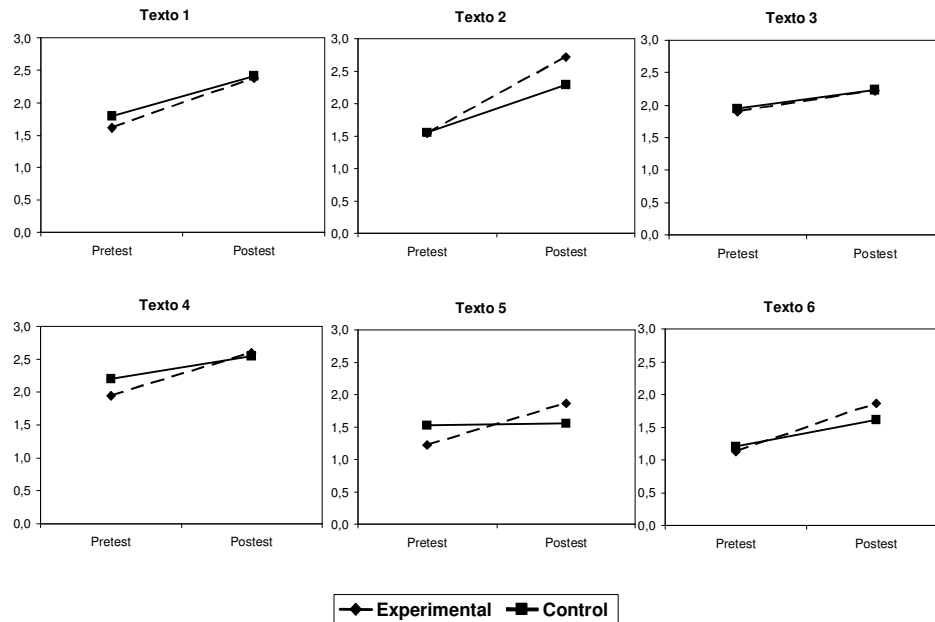
<p><i>Tabla 4</i> Diferencias significativas (Scheffé) entre el pretest, posttest y retest para el proceso de planificación, en los diferentes textos y para los tres grupos (grupo experimental, grupo control y buenos escritores)</p>					
Texto	Evaluación		Grupo experimental	Grupo control	Buenos escritores
1. Tema libre	pretest	posttest	,003**	,050*	,726
	pretest	retest	,003**	,198	,327
	posttest	retest	1,00	,517	,528
2. Cuento con personaje principal determinado	pretest	posttest	,000***	,005**	,009**
	pretest	retest	,284	,555	,659
	posttest	retest	,000***	,022*	,003**
3. Cuento popular	pretest	posttest	,142	,297	,202
	pretest	retest	,293	,916	,379
	posttest	retest	,673	,347	,689
4. Describir un dibujo	pretest	posttest	,017*	,188	,044*
	pretest	retest	,196	1,00	,331
	posttest	retest	,261	,188	,290
5. Describir un lugar	pretest	posttest	,004**	,902	,710
	pretest	retest	,028*	,180	,414
	posttest	retest	,457	,222	,236
6. Comentario sobre situación concreta	pretest	posttest	,002**	,043*	,142
	pretest	retest	,000***	,000***	,016*
	posttest	retest	,612	,081	,326
*p<.05; **p<.01; ***<.001					

El comentario que se hace de la tabla ejemplifica la inconveniencia del procedimiento analítico utilizado:

“En el proceso de *planificación* (ver Tabla 4) encontramos que *para el grupo experimental hay diferencias significativas entre el pretest y el posttest en los textos 1, 2, 4, 5 y 6, diferencias que se mantiene entre el pretest y el retest para los textos 1, 5 y 6. En el grupo control se observan diferencias significativas en textos 1, 2 y 6 y se mantienen para el texto 6...*” (cursiva mía)

Este párrafo ilustra perfectamente por qué es necesario desarrollar los experimentos con grupo control, ya que si tanto en el grupo experimental como en el grupo control se observan diferencias entre pretest y posttest en los textos 1, 2 y 6, es posible que las diferencias encontradas en el grupo experimental no sean la consecuencia de la intervención realizada, sino de alguna otra circunstancia, por ejemplo, la maduración de los sujetos, que afecta tanto a uno como a otro grupo. Sólo si el grupo experimental evoluciona de una forma diferente que el grupo control, lo que eventualmente solo podrá ser puesto de manifiesto por un análisis de interacción, se podría concluir que en el grupo experimental sucede algo diferente a lo que sucede en el grupo control.

En la Figura 4.12 se muestra una representación gráfica preparada a partir de los resultados descriptivos presentados en la Tabla 2 del artículo (los que se analizan en la Tabla 4) para el proceso de planificación en los grupos experimental y control en pretest y posttest (los únicos que se ofrecen en el artículo, pues en la Tabla 2 no se presentan las descriptivas para el grupo de alumnos sin DAE). A la vista de los gráficos los resultados de ambos grupos son tan parecidos en la mayoría de los textos que resulta difícil aceptar la conclusión presentada por los autores de que “*aunque el grupo control mejoró en los procesos de escritura entre el pretest y el posttest, los estudiantes con DA mejoraron en una proporción más acelerada.*” (cursiva mía).

**Figura 4.12. Representación gráfica a partir de la Tabla 2 del original (*Psicothema*, 2003)**

**Ejemplo 15) (IJCHP, 2005)** estudiaron la eficacia de un tratamiento cognitivo-conductual para la preocupación excesiva en personas mayores por medio de un estudio pre-post con grupo control. Sin embargo, para analizar los resultados no se utilizó un ANOVA de dos factores, un factor intrasujetos, *momento* (con dos niveles: pre; post), y otro intersujetos, *grupo* (con dos niveles: clínico; control), para centrarse en el análisis de interacción. Por el contrario, se utilizó la prueba T para muestras relacionadas en cada uno de los *grupos* y la prueba T para muestras independientes en cada uno de los *momentos*. Los autores describen el análisis realizado de la siguiente manera (pág. 220)

“En primer lugar se realizó una prueba *t* de comparación de medias entre el grupo experimental y el control sobre las puntuaciones obtenidas previamente a la aplicación del programa. No se encontraron diferencias estadísticamente significativas entre grupos en ninguna de las variables evaluadas. En la Tabla 1 pueden verse la media y la desviación típica de las puntuaciones de los dos grupos en las diferentes variables evaluadas antes como después de la aplicación del programa.

Tras el tratamiento se compararon de nuevo las puntuaciones entre los grupos mediante la prueba  $t$  para muestras independientes. El grupo clínico presenta puntuaciones significativamente inferiores en preocupación-rasgo ( $t(39) = 2,57$ ;  $p < 0,05$ ) y severidad de la ansiedad generalizada ( $t(39) = 2,04$ ;  $p < 0,01$ ), pero no en metapreocupación respecto al grupo control. La variación en las puntuaciones entre las dos aplicaciones de la evaluación fue analizada mediante pruebas  $t$  para muestras relacionadas. En el grupo control no se produjeron diferencias estadísticamente significativas entre la evaluación previa al tratamiento y la evaluación posterior. En el grupo clínico en cambio, las puntuaciones se redujeron significativamente entre una y otra en las escalas para la preocupación-rasgo ( $t(16) = 8,98$ ;  $p < 0,01$ ), en severidad de la ansiedad generalizada ( $t(16) = 7,40$ ;  $p < 0,01$ ) y en metapreocupaciones ( $t = 4,47$ ;  $p < 0,01$ )”

En la Tabla 4.6 se reproduce la Tabla 1 del artículo de referencia. Tanto en el resumen del artículo como en sus conclusiones se reivindica la eficacia del tratamiento para el afrontamiento clínico de la preocupación excesiva en la vejez:

“...el efecto pre-post y el porcentaje de cambio fueron elevados para el grupo experimental y prácticamente nulos en el grupo control, de modo que *el programa realizado parece ser eficaz en la reducción de la preocupación excesiva en personas mayores*” (cursiva mía).

La conclusión de eficacia (comparativa con el grupo control) implica la existencia de interacción que excede de lo demostrado, ya que en ningún momento se comparan formalmente los efectos pre-post entre ambos grupos o las diferencias clínico-control entre los momentos pre-post. En este caso, el análisis de los efectos simples lleva seguramente a la misma conclusión que se habría obtenido a partir de un análisis de interacción correctamente desarrollado, pero el análisis de los efectos simples no es informativo con respecto a la interacción y además puede resultar engañoso.

Tabla 4.6. Tomada del original (*IJCHP*, 2005)

TABLA 1. Puntuaciones medias (y desviaciones típicas) para cada una de las condiciones de tratamiento y momento de evaluación.

	<i>Grupo clínico</i>		<i>Grupo control</i>	
	<i>Pre</i>	<i>Post</i>	<i>Pre</i>	<i>Post</i>
Preocupación-rasgo	30,8 (7,3)	23,3 (5,6)	28,3 (9,0)	29,4 (9,6)
Metapreocupación	12,8 (4,6)	10,5 (3,5)	11,6 (3,9)	11,5 (3,8)
Severidad de la ansiedad generalizada	39,8 (11,3)	27,9 (8,8)	35,0 (16,4)	35,8 (16,0)

Se podría argumentar que el análisis de los efectos simples para interpretar la interacción solo es lícito cuando anteriormente se ha demostrado mediante la prueba  $F$  correspondiente que existe una interacción significativa. Sin embargo, el resultado del análisis de los efectos simples no está definido en función del resultado del análisis de la interacción, ya que en los efectos simples intervienen tanto los efectos principales como la interacción, dando lugar a diferentes situaciones y haciendo que el análisis de los efectos simples no facilite una explicación adecuada de la interacción por más que en algunos casos pueda llevar a las mismas conclusiones. Como ya se ha comentado, esta interpretación, que está extendida incluso en ámbitos académicos (Zuckerman, 1993), no sólo es incorrecta en sí misma, sino que presta indirectamente apoyo a la aparente inutilidad del análisis de la interacción como “paso intermedio” de escaso interés y tiende a justificar el tipo de error que se acaba de presentar.

A continuación se presentan ejemplos en que el análisis de los efectos simples es utilizado como explicación de la interacción entre dos factores.

**Ejemplo 16)** (*Psicothema*, 2005) Los autores estudiaron cómo afecta la inclusividad al favoritismo endogrupal y la deshumanización del exogrupo, midiendo las actitudes positivas hacia el exogrupo y el endogrupo. Para ello utilizaron un diseño de cuatro

factores: un factor intersujetos, *identidad activada* (con tres niveles: supranacional; nacional; regional), y tres factores intrasujetos, *grupo estímulo* (con dos niveles: endogrupo; exogrupo), *tipo de emoción* (con dos niveles: sentimientos; emociones) y *valencia* (con dos niveles: positiva; negativa). Para estudiar el efecto de la inclusividad llevaron a cabo un ANOVA de dos factores, *identidad activada*  $\times$  *valencia*, para determinar si el nivel de inclusividad afectaba al número de rasgos positivos y negativos. En el apartado de resultados se dice (pág. 444):

“Los resultados muestran sólo un efecto principal de la variable valencia con más atribución de los rasgos negativos ( $M = 4,12$ ) que positivos ( $M = 3,05$ ;  $F_{(1, 65)} = 14,6$ ;  $p < .001$ ). Este efecto queda matizado por la significación derivada de la interacción entre valencia e inclusividad ( $F_{(2, 65)} = 8,52$ ;  $p < .001$ ) ya que rasgos positivos y negativos siguen una tendencia inversa en relación con los niveles de inclusividad. *El análisis de los efectos simples de la interacción reveló que la significación se debía a la diferencia entre el nivel de activación regional y el nivel de activación supranacional, tanto respecto a los rasgos positivos ( $M = 2,48$  y  $3,60$ , respectivamente,  $F_{(1, 65)} = 7,83$ ;  $p = .007$ ) como a los rasgos negativos ( $M = 4,96$  y  $3,35$ , respectivamente,  $F_{(1, 65)} = 11,30$ ;  $p < .001$ ). Los datos relativos a la activación nacional (España) no muestran diferencias significativas frente a los otros niveles de activación, ni en los rasgos positivos ( $M = 3,05$ ) ni en los rasgos negativos ( $M = 4,05$ ), aunque contribuyen a reflejar la tendencia defendida en la hipótesis.” (cursiva mía).*

Lo que se está diciendo en este párrafo es que la interacción es la consecuencia de los efectos simples que alcanzaron significación estadística, considerando por separado los efectos simples que no alcanzaron significación estadística. En realidad la interacción se debe a que la diferencia entre las atribuciones positivas en los niveles de inclusividad supranacional y regional difiere significativamente de la diferencia entre las atribuciones negativas en dichos niveles de inclusividad. La posible contribución de otros componentes en la interacción requiere de un análisis pormenorizado no presentado en



el artículo. Esta misma forma de abordar la interacción se presenta más adelante ilustrando los resultados con la Figura 1 del artículo, que se reproduce en la Figura 4.13.

En las págs. 444-5 se señala:

“...Los resultados muestran un efecto principal de la variable valencia que queda matizado por la interacción significativa de las dos variables ( $F_{(2, 65)}=6,12$ ;  $p=.004$ ).

Como se observa en la Figura 1, la tendencia a infrahumanizar se ve afectada por el fuerte papel que desempeña la valencia de los sentimientos. El análisis de los efectos simples de la interacción deja constancia de los contrastes significativos. Concretamente se encuentran diferencias entre el nivel regional y supranacional en los sentimientos positivos ( $M= 1,4$  y  $0,68$ , repectivamente;  $F_{(1, 65)}=6,93$ ;  $p=.01$ ) y en los sentimientos negativos ( $M= 1,5$  y  $2,2$ , repectivamente;  $F_{(1, 65)}=5,1$ ;  $p=.027$ ). Asimismo, se ha encontrado un efecto simple de la interacción significativa entre el nivel regional y el nacional en los sentimientos negativos ( $M= 1,5$  y  $2,2$ , repectivamente;  $F_{(1, 65)}=4,04$ ;  $p=.05$ ).

El procedimiento utilizado por los autores para explicar la interacción encontrada es analizar los efectos simples, determinando los pares de niveles que alcanzaron significación estadística. Mientras que a partir del gráfico cabe suponer que la interacción *identidad activada*  $\times$  *valencia* involucra diferencias significativas entre diferencias de sentimientos en los niveles regional y supranacional, el análisis realizado no permite saber si se producen diferencias significativas entre los niveles regional y nacional, lo que aparentemente es posible teniendo en consideración el gráfico presentado por los autores.

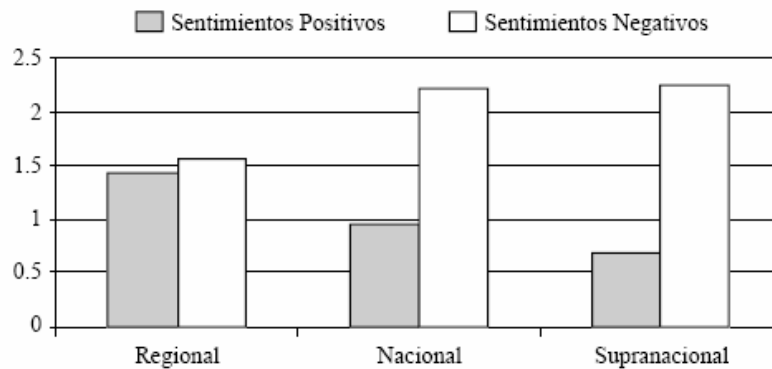
Figura 4.13. Tomada del original (*Psicothema*, 2005)


Figura 1. Asignación de sentimientos positivos y negativos al exogrupo «magrebies» en función del nivel de inclusividad de la identidad activada

**Ejemplo 17)** (*Psicothema*, 2002) Los autores estudiaron la influencia del tipo de recuerdo y el intervalo de retención sobre diversos aspectos del recuerdo. Se trata de un diseño con dos factores intrasujeto: *tipo de recuerdo* (con tres niveles: autobiográfico; noticias; conceptos académicos) y *periodo temporal del recuerdo* (con dos niveles: esta semana; un mes atrás). Se estudia la influencia sobre tres variables dependientes: número de recuerdos; detalles perceptivos y esfuerzo cognitivo, utilizando un ANOVA factorial intrasujeto. En los tres análisis se obtiene una interacción significativa, administrándose a continuación una “*prueba de los efectos simples para el análisis de la interacción*”, consistente en determinar los efectos simples de *periodo temporal* en cada nivel de *tipo de recuerdo*. La Figura 1 del artículo, que se reproduce en la Figura 4.14, se utiliza para discutir los resultados obtenidos. En el apartado de resultados (pág. 395) se señala:

Sobre el número de recuerdos: “...Una prueba de los efectos simples para el análisis de la interacción (ver panel superior de la Figura 1) determinó como significativos

*( $p < 0,001$ ) los efectos del tiempo tanto sobre los recuerdos personales como sobre las noticias, pero no así sobre los conceptos” (cursiva mía).*

La interacción se debe a que la diferencia entre “la misma semana” y “un mes atrás” es significativamente mayor en “noticias” que en “conceptos académicos”. También, posiblemente, a que la diferencia entre el número de recuerdos personales recordados en la misma semana y un mes atrás es significativamente mayor que la diferencia entre el número de conceptos académicos recordados en la misma semana y un mes atrás, pero para conocer dicho extremo sería necesario desglosar la interacción por medio de comparaciones de diferencias entre niveles.

Sobre los detalles perceptivos (pág. 396): “...Por último, una prueba de los efectos simples para el análisis de la interacción (ver panel central de la Figura 1) determinó como significativos ( $p < 0,001$ ) los efectos del tiempo sobre las tres condiciones de recuerdos” (cursiva mía).

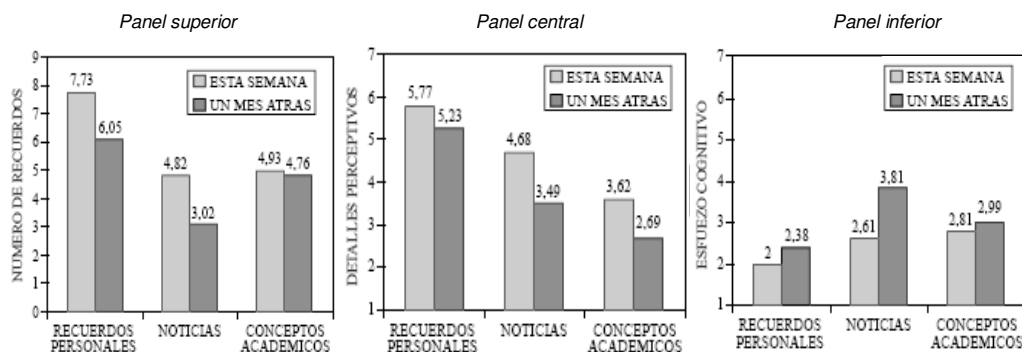
Esta descripción no tiene ninguna relación con la interacción, pues en ninguna manera ilustra en qué modo el efecto sobre la dependiente de un factor está mediado por los niveles del otro factor, ya que si todos los efectos simples de *tiempo* son significativos, independientemente del nivel de *recuerdo* considerado, lo que se está describiendo es un efecto principal, y no se puede vislumbrar en que forma se produce la interacción encontrada. El texto elegido muestra la incapacidad del abordaje basado en los efectos simples para explicar una interacción significativa. La interacción se debe a que la diferencia entre los detalles perceptivos de “recuerdos personales” recordados en “la misma semana” y “un mes atrás” es significativamente menor que la diferencia entre los detalles perceptivos de “noticias” recordados “en la misma semana” y “un mes atrás”. De nuevo es posible que la interacción se deba además a que las diferencias entre los detalles perceptivos en “la misma semana” y “un mes atrás” sean distintas para

“noticias” y “conceptos académicos”, pero para aclarar este extremo sería necesario desmenuzar la interacción en sus componentes.

Sobre el esfuerzo cognitivo (pág. 396): “...Por último, una prueba de los efectos simples para el análisis de la interacción (ver panel inferior de la Figura 1) determinó como significativos ( $p < 0,001$ ) los efectos del tiempo tanto sobre los recuerdos personales como sobre las noticias, pero no así sobre los conceptos” (cursiva mía).

El comentario para este párrafo es similar al realizado cuando la variable dependiente es el número de recuerdos. La interacción se debe a que la diferencia entre “la misma semana” y “un mes atrás” es significativamente mayor en “noticias” que en “conceptos académicos”.

**Figura 4.14. Tomada del original (*Psicothema*, 2002)**



*Figura 1. Promedios del número de recuerdos elicitados (panel superior), detalles perceptivos (panel central) y esfuerzo cognitivo (panel inferior) para las seis condiciones del estudio*

**Ejemplo 18) (*Psicothema*, 2002)** Los autores estudiaron la reactividad cardiovascular en sujetos con *Locus* de control (LOC) interno y externo durante la realización de una tarea física presentada con o sin información sobre el rendimiento durante el desarrollo. Se trata de un diseño con dos factores: uno intrasujetos, *condición* (con dos niveles: sin información; con información), y otro intersujetos, *Locus de control* (con dos niveles:

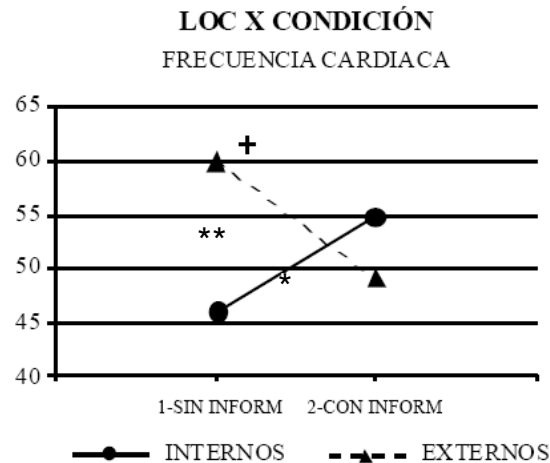
interno; externo). Los Figura 2 del estudio se reproduce en la Figura 4.15. En el apartado resultados (pág. 586) se dice:

“Por una parte, en el análisis de la reactividad de FC en *tarea*, se obtuvo una interacción significativa LOC  $\times$  Condición [ $F_{(1, 28)} = 8,03, p < 0,01$ ](ver Figura 2), *indicando que los sujetos internos presentaban una mayor reactividad en la condición con información que en la condición control* [ $t_{(14)} = 2,18, p < 0,05$ ]; *mientras que los sujetos externos tendían a presentar mayor reactividad en la condición 1 o sin información que en la 2 o con feedback* [ $t_{(14)} = 1,94, p < 0,07$ ]. Asimismo, *había diferencia entre ambos grupos en la condición 1 o sin información, de forma que los externos se reactivaban significativamente más que los internos* [ $F_{(1, 29)} = 6,003, p < 0,02$ ](cursiva mía).

En este párrafo se confunde el efecto de la interacción con el de los efectos simples, “*la interacción indica que... existen efectos simples significativos*”, cuando son hallazgos que no están directamente relacionados, pues puede haber interacción en presencia de efectos simples todos ellos significativos, en presencia de efectos simples todos ellos no significativos, o en presencia de efectos simples unos significativos y otros no significativos. Por otra parte, en tanto que se trata de un diseño  $2 \times 2$ , tan sólo se dispone de un grado de libertad, por lo que la interpretación de la interacción no requiere ningún análisis adicional. Bastaría con haber interpretado correctamente la interacción sobre la tabla o gráfico (la misma que se presenta pero sin necesidad de analizar los efectos simples) de la siguiente manera: “la diferencia en la reactividad entre los sujetos con LOC interno y externo cuando realizan la tarea sin información es significativamente diferente, y de sentido contrario, de la diferencia en la reactividad entre los sujetos con LOC interno y externo cuando realizan la tarea con información”. Igualmente, la interacción podría haberse explicado de la siguiente manera: “la diferencia en la reactividad de los sujetos con LOC interno cuando realizan la tarea sin información y con información es significativamente diferente, y de sentido contrario, de la diferencia

en la reactividad de los sujetos con LOC externo cuando realizan la tarea sin información y con información” y para la que tampoco habría sido necesario ningún contraste adicional.

**Figura 4.15. Tomada del original (Psicothema, 2002)**



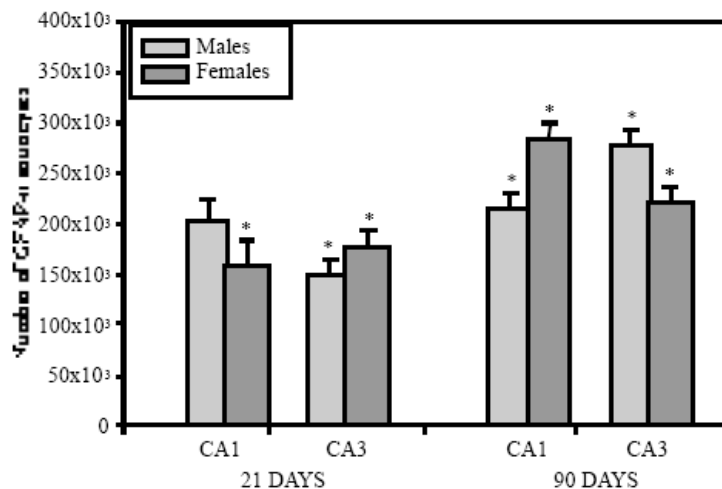
*Figura 2. Interacción entre la variable de «Locus de control» y el factor «Condición» en el primer estudio para la reactividad de frecuencia cardíaca en Tarea (restando el valor de línea base a la puntuación máxima en cada tarea) [Notas: +  $p < .07$ ; \*  $p < .05$ ; \*\*  $p < .02$ ]*

**Ejemplo 19) (Psicothema, 2003)** Los autores estudiaron el desarrollo postnatal de la astrogliá en función del sexo. Para ello contaron el número de astrocitos en las áreas CA1 y CA3 del hipocampo de ratas machos y hembras de 21 y 90 días de edad. Se trata de un diseño con 3 factores: 2 factores intersujetos, *sexo* (con dos niveles: macho; hembra) y *edad* (con dos niveles: 21 d.; 90 d.) y un factor intrasujetos, *área* (con dos niveles: CA1; CA3). En la Figura 4.16 se reproduce la Figura 2 del artículo, presentando el número medio de astrocitos inmunoreactivos a la proteína ácida de la glía en las áreas CA1 y CA3 de ratones machos y hembras sacrificados a los 21 y 90 días de desarrollo postnatal. En el apartado de resultados, la interacción *género*  $\times$  *área* se describe como sigue (pág. 218):

“...However a significant interaction was found between gender and hippocampal area in adult (90 days) rats [ $F(1,20) = 14,96$ ,  $p < 0,001$ ]. *Particularly, the number of GFAP-ir astrocytes was significantly higher in the CA3 area of males as compared with age-matches females [Tukey’s test;  $p < 0,05$ ]. Moreover, females showed significantly more GFAP-ir astrocytes in the CA1 area as compared to males [Tukey’s test;  $p < 0,05$ ].* In addition, the magnitude of the differences between the hippocampal areas studied was similar in males and females (mean numerical differences of about 22%) (cursiva mía)

La interacción se debe a que la diferencia en el número de astrocitos en machos y hembras difiere significativamente entre CA1 y CA3. La interacción no depende de la significación de los efectos simples ni de su magnitud, igual en ambos casos en el ejemplo, sino de la diferencia entre ellos.

**Figura 4.16. Tomada del original (Psicothema, 2003)**



**Figure 2.** Total number (mean + S.E.M.) of GFAP-ir astrocytes quantified in the hippocampus (CA1 and CA3 areas) of male and female rats from both sexes. \*  $p < 0.05$  males versus females, and +  $p < 0.05$ , comparison between 21 and 90 days of age (Tukey’s post-hoc tests)

En los siguientes ejemplos se presentaran una selección de artículos en los que la falta de capacidad explicativa de los efectos simples con respecto a la interacción queda en evidencia, en algunos casos incluso para los propios autores de los artículos.

**Ejemplo 20)** (*Psicothema*, 2002) analizó las diferencias de género en las explicaciones causales de los adolescentes en tareas resolubles e irresolubles. Se trata de un diseño con dos factores intersujetos, *sexo* (con dos niveles: hombre; mujer) y *tarea* (con dos niveles: resoluble; irresoluble). Existen cuatro tipos de explicaciones causales o variables dependientes (dificultad; esfuerzo; suerte; habilidad). En la Figura 1 del artículo, que se reproduce en la Figura 4.17, se presentan los resultados, donde es de resaltar que unas interacciones son significativas y otras no, lo que no guarda ninguna relación con que los efectos simples alcancen o no significación estadística. En la pág. 574 se indica:

“La interacción entre el género y el tipo de tarea no resultó significativa cuando se tuvieron en cuenta las explicaciones causales de “dificultad”,  $F(1, 68) = 0,26$ ,  $p > 0,05$  y “esfuerzo”  $F(1, 68) = 3,4$ ,  $p = 0,07$  (Figuras 1.A y 1.B). En este último caso los análisis indicaron diferencias de género solo en la tarea resoluble,  $t(32) = -3,01$ ,  $p < 0,01$ . En concreto, las mujeres explicaron que su rendimiento se debía al esfuerzo ( $M = 5,12$ ,  $DS = 1,82$ ) en mayor medida que los hombres ( $M = 3,44$ ,  $DS = 1,42$ ).

En la Figura 1.A (dificultad) no se encuentra interacción significativa y los efectos simples de *sexo* no son significativos en ninguno de los dos tipos de *tarea*. En la Figura 1.B (esfuerzo) la interacción tampoco es significativa, sin embargo, uno de los efectos simples sí lo es y el otro no (véase la Figura 4.17). En la misma página, con respecto a la variable dependiente “suerte” se indica:

“Se halló una interacción significativa en la explicación causal de “suerte”  $F(1, 68) = 4,64$ ,  $p < 0,05$  (fig. 1.C.). Las mujeres del grupo irresoluble ( $M = 3,75$ ) estimaron que habían tenido menos suerte que las del grupo resoluble ( $M = 7$ ),  $t(30) = -3,7$ ,  $p < 0,01$ . Los hombres, en cambio, no mostraron diferencias ( $M$ s 5,44 y 6,06)  $t(30) = -0,72$ ,  $p > 0,05$ .



En la Figura 1.C (véase la Figura 4.17) si hay interacción *sexo*  $\times$  *tarea*. Para explicarla no se recurre a los efectos simples de *sexo* en cada nivel de *tarea*, sino a los efectos simples de *tarea* en cada nivel de *sexo* (un efecto simple significativo y el otro no). Por otra parte, no se facilitan los datos necesarios para poder calcular los efectos simples de *sexo* en cada nivel de *tarea*, aunque a juzgar por el gráfico es posible que ninguno de ellos sea significativo (quizá por esta razón se elude interpretar la interacción mediante los efectos simples de la variable *sexo*).

**Figura 4.17. Tomada del original (*Psicothema*, 2002)**

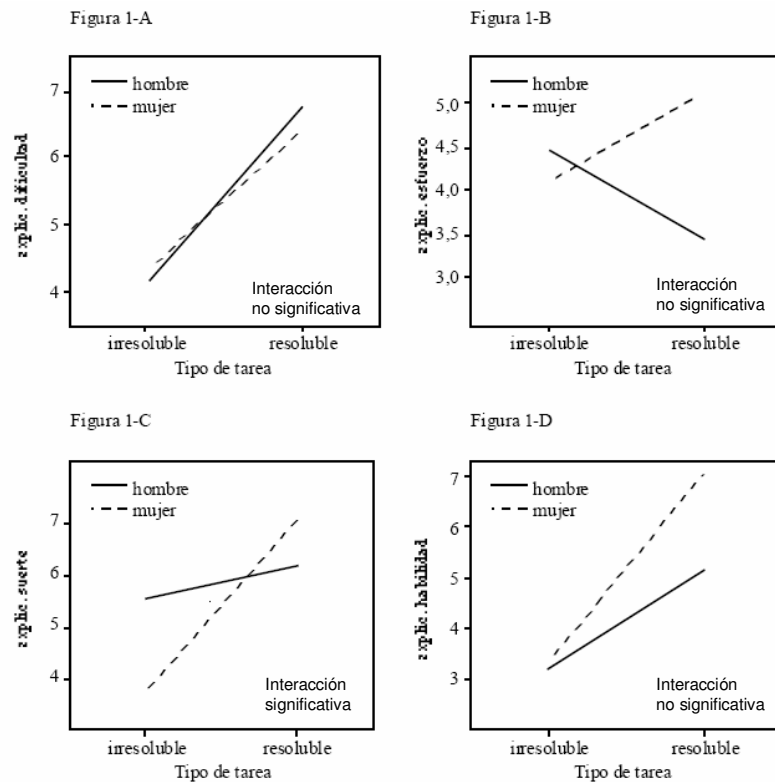


Figura 1. Puntuaciones medias de los grupos en las explicaciones causales de dificultad (1-A), esfuerzo (1-B), suerte (1-C) y habilidad (1-D)

Finalmente, para la dependiente dependiente “habilidad” se indica:

“No hubo interacción significativa entre el genero y el tipo de tarea en las explicaciones a la “habilidad”,  $F(1, 68) = 2,39$ ,  $p > 0,05$  (fig. 1.D). Si bien no hubo diferencias entre hombres ( $M = 3,17$ ,  $DS = 2,20$ ) y mujeres ( $M = 3,38$ ,  $DS = 2,16$ ) en la condición

irresoluble,  $t(32) = -0,27$ ,  $p > 0,05$ , en la condición resoluble las mujeres ( $M = 7,13$ ,  $DS = 2,25$ ) explicaron su rendimiento en términos de una mayor habilidad que los hombres ( $M = 5,11$ ,  $DS = 2,89$ ),  $t(32) = -2,25$ ,  $p < 0,05$ .

Se vuelve a la línea de explicación inicial a pesar de haber tenido que eludirla en la explicación de la interacción significativa referida a la “suerte”.

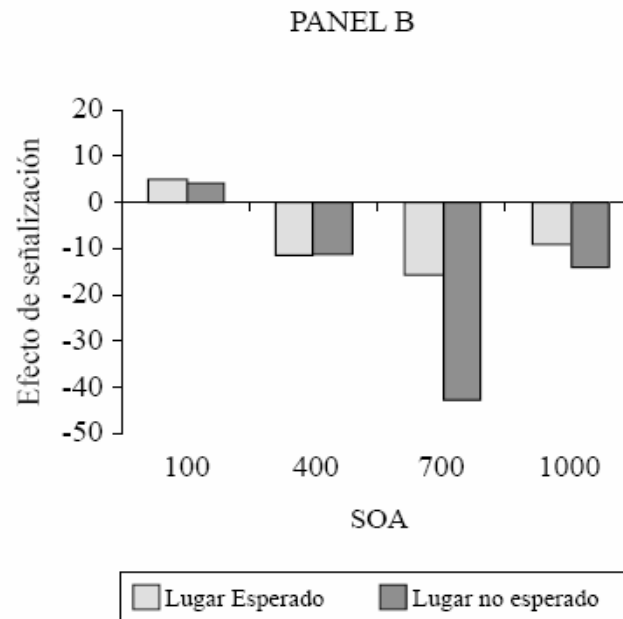
**Ejemplo 21** (*Psicothema*, 2004) Los autores estudiaron la inhibición de retorno (IR) en términos de incremento en el tiempo de reacción en lugares a los que se ha orientado la atención anteriormente. Estudiaron el tiempo de reacción en dos tipos de tareas, detección y discriminación, manipulando la aparición o no de una señal previa al objetivo (señalización), el tiempo transcurrido desde que aparece la señal hasta que aparece el objetivo (Stimulus Onset Asynchrony, SOA) y la aparición de la señal en el mismo o diferente campo visual que el objetivo (expectativa espacial). Se trata de un diseño con cuatro factores: tres factores intrasujetos, *señalización* (con dos niveles: señalado; no señalado), *SOA* (con cuatro niveles: 100 ms; 400 ms; 700 ms; 1000 ms) y *expectativa espacial* (con dos niveles: esperado; no esperado) y un factor intersujetos, *tarea* (con dos niveles: detección; discriminación). En el apartado de resultados se estudia la posible interacción *tarea*  $\times$  *expectativa espacial*  $\times$  *señalización* para cada nivel de *SOA* y, dentro de cada nivel, las interacciones de primer orden significativas. En general, las interacciones de primer orden que alcanzan significación estadística se explican en términos de efectos simples, uno estadísticamente significativo y otro no. Sin embargo, en el apartado de resultados en referencia al *SOA 700* se dice (pág. 252):

*SOA 700: ... En este caso, la interacción Expectativa Espacial  $\times$  Señalización también resultó significativa,  $F(1, 29) = 25,32$ ,  $p < 0,0005$ . Este dato podría estar, en principio, a favor de la hipótesis de la orientación-reorientación de la atención. Sin embargo,*

*hemos de destacar que, a pesar de la interacción significativa, el efecto de IR fue altamente significativo tanto en el lugar esperado como en el no esperado ( $p < 0,00001$  en ambos casos) (cursiva mía)*

El hallazgo de interacción en el SOA 700 sí resulta acorde con la hipótesis de orientación-reorientación pues se debe interpretar de la siguiente manera: la diferencia entre los tiempos de reacción entre los objetivos señalados y no señalados es mayor cuando éstos no son esperados que cuando son esperados. Esta interpretación es acorde con la representación gráfica de los resultados que aparece en la Figura 4.18 que reproduce el panel B de la Figura 1 del artículo.<sup>7</sup>

**Figura 4.18. Tomada del original (*Psicothema*, 2004)**



**Figura 1.**

*En el panel B podemos observar el efecto de señalización en los diferentes SOAs en función de la Expectativa Espacial (Lugar Esperado vs. No Esperado). En los ensayos esperados el estímulo objetivo se presenta en el lugar donde el participante espera (por la información del punto de fijación), mientras que en los no esperados el estímulo objetivo se presenta en el lugar no esperado por el participante*

*Efecto de señalización: TR en los ensayos no señalados menos TR en los ensayos señalados*

<sup>7</sup> Para interpretar correctamente la Figura 4.18 hay que tener en cuenta que cada barra del diagrama se refiere a la inhibición de retorno, IR, diferencia entre ensayos señalizados y no señalizados. En el SOA700 la IR es significatva tanto cuando el estímulo objetivo se presenta en el lugar esperado como en el no esperado.

Sin embargo, la interpretación que se hace a partir de los efectos simples no arroja luz sobre el sentido de la interacción por lo que los autores dudan sobre si facilita o no apoyo a la hipótesis de orientación-reorientación, para acabar llamando la atención sobre el hallazgo “contradictorio” de una interacción significativa que ampara dos efectos simples “altamente significativos”.

**Ejemplo 22)** (SJP, 2001) Los autores estudiaron la influencia de los factores estresantes en el funcionamiento cognitivo. Cada miembro de un grupo de niños de 5 a 12 años de un barrio de bajo nivel socioeconómico fue clasificado en función de haberse sentido afectado por disparos y peleas. Todos los niños fueron expuestos a dos estímulos visuales diferentes, el primero consistente en 63 imágenes de un encuentro amistoso entre dos vaqueros, en las que paulatinamente se van perdiendo detalles (LSFT) y, el segundo, en 63 imágenes de un tiroteo entre dos vaqueros, en las que paulatinamente se van perdiendo detalles (LSSOT). Entre otras dependientes se midió el número de cambios detectados (correctamente) y el momento en que se detecta el cambio, construyendo a partir de ambas dependientes un índice que es menor cuanto mayor y más rápido es el número de cambios detectados y mayor cuando menor y más lento se detectan dichos cambios. También se midió el número de detalles recordados de la figura del amigo en el episodio LSFT y de la víctima del tiroteo en el episodio LSSOT. Se trata de un diseño con dos factores: uno intrasujeto, *estímulo* (con dos niveles: LSFT; LSSOT) y otro intersujetos, *afectación por disparos/peleas*, (con dos niveles: sí; no). Los resultados sobre dos dependientes diferentes, escenas recordadas y detalles de las figuras humanas recordados se presentan en la Figura 2 del artículo, que se reproduce en la Figura 4.19. En el apartado de resultados (pág. 43), se señala:

“ANOVA results showed a statistically significant interaction effect  $F(1, 91) = 5,78$ ,  $p = 0,01$ , size effect 0,33, power = 0,707, for shootings/fights on LSSOT versus LSFT in correct change ratio (see figure 2, left graph). Children who reported shooting/fights detected fewer changes and later (cognitive levelling) with the shoot-out scene (LSSOT correct ratio  $M = 23.5$ ,  $SD = 4.1$ ) versus the friendship scene (LSFT correct ratio  $M = 22.2$ ,  $SD = 4.6$ ). In contrast, children who did not report the stressor detected more changes and sooner (cognitive sharpening) with the shoot-out scene (LSSOT correct ratio  $M = 21.8$ ,  $SD = 3.5$ ) versus the friendship scene (LSFT correct ratio  $M = 23.7$ ,  $SD = 4.1$ ). *The difference between the groups in remembering details of the shoot-out scene was not statistically significant,  $t(91) = 1,37$ ,  $p = 0,12$ . Nor was the difference in the remembering details of the friendship scene  $t(91) = 1,14$ ,  $p = 0,15$ . However, as noted the interaction between test stimuli and reporting shootings/fights was statistically significant....*

ANOVA results showed a statistically significant interaction effect  $F(1, 91) = 13,18$ ,  $p = 0,001$ , size effect 0,33, power = 0,89, for shootings/fights on LSSOT victim versus LSFT friend scores. As shown in Figure 2, right graph... [a continuación se pasa a describir los efectos en cada una de las casillas] *The difference between groups in the number of changes perceived on the person of the victim was not statistically significant,  $t(91) = 1,04$ ,  $p = 0,20$ . The difference between groups in the number of changes perceived on the person of the friend was statistically significant,  $t(91) = 2,69$ ,  $p = 0,001$ .*

En el segundo párrafo, la interacción resulta estadísticamente significativa y para analizarla es desglosada en los efectos simples, encontrando que uno de ellos es significativo y el otro no. Como este hallazgo se corresponde con una forma muy habitual de entender la interacción entre dos factores, no llama la atención a los autores. Sin embargo, en el primer párrafo en el que se procedió de igual manera, es decir al encontrar una interacción significativa se analizaron los efectos simples, al encontrar que ninguno de ellos es significativo el investigador añade el comentario: “*However, as noted the interaction between test stimuli and reporting shootings/fights was statistically significant.*”

Figura 4.19. Tomada del original (SJP, 2001)

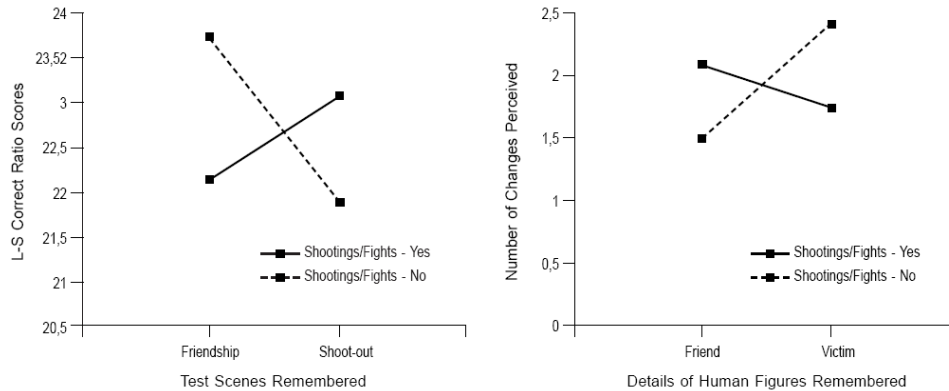


Figure 2. Reporting shootings/fights and remembering details of pictures depicting a friendship and shoot-out and details on the persons of a friend and a victim (LSFT vs. LSSOT).

En la Figura 4.19 (izquierda) se observa que los efectos simples tienen sentido contrario y que por tanto en el análisis de la interacción, considerado como diferencia entre diferencias, sus magnitudes se sumarán, haciendo posible que la interacción sea significativa aunque ninguno de los efectos simples lo sea por sí mismo. De nuevo queda patente la necesidad de que los investigadores tomen conciencia de que para interpretar una interacción en un ANOVA 2x2 no es imprescindible hacer ningún contraste adicional a la  $F$  de la interacción y que la significación de los efectos simples no está predefinida por la significación de la interacción.

**Ejemplo 23)** (Psicothema, 2001) Los autores estudiaron el número de palabras recordadas por sujetos depresivos y un grupo control en función del tipo de palabra (adjetivos relacionados con la personalidad). Se trata de un diseño con dos factores: un factor intrasujetos, *tipo de palabra* (con tres niveles: negativo-depresivo; positivo-depresivo; positivo-control) y un factor intersujetos, *grupo* (con dos niveles: depresivos; control). Los resultados se presentan en la Tabla 2 del artículo referenciado que se reproduce en la Tabla 4.7. En el apartado resultados (pág. 195) se dice:

“... Asimismo, el análisis de los datos también mostró la existencia de un efecto significativo para la interacción de primer orden Grupo  $\times$  Tipo de palabras,  $F(2, 56) = 16,06$ ,  $p < 0,001$ .

*El análisis a posteriori de la interacción Grupo  $\times$  Tipo de palabras encontró un efecto significativo en el factor Grupo sobre el recuerdo de palabras negativas-depresivas,  $F(1, 28) = 9,83$ ,  $p < 0,01$ , positivas-depresivas,  $F(1, 28) = 105,09$ ,  $p < 0,0001$ , positivas-controles,  $F(1, 28) = 29,42$ ,  $p < 0,001$ . Los sujetos del grupo control recordaron mas palabras (negativas-depresivas, positivas-depresivas y positivas-controles) que los sujetos depresivos.*

Igualmente se encontró un efecto significativo del Tipo de palabras para los sujetos depresivos,  $F(2, 28) = 7,77$ ,  $p < 0,01$ . Las comparaciones a posteriori mediante la prueba de Newman-Kelus mostraban que los sujetos depresivos recordaban más palabras negativas-depresivas ( $X = 1,33$ ) que positivas-depresivas ( $X = 0,37$ ) y positivas-controles ( $X = 0,33$ )( $p < 0,001$ ).

También apareció un efecto significativo del factor tipo de palabras en los sujetos no depresivos,  $F(2, 28) = 10,66$ ,  $p < 0,001$ . Las comparaciones a posteriori mediante la prueba de Newman-Keuls mostraron que los sujetos no depresivos recordaban más palabras positivas-depresivas ( $X = 5,87$ ) que negativas-depresivas ( $X = 3,00$ ) ( $p < 0,01$ ) y postivas.-controles ( $X = 4,12$ ) ( $p < 0,05$ ). Asimismo, los sujetos no depresivos recordaron más palabras positivas-controles ( $X = 4,13$ ) que negativas-depresivas ( $X = 3,00$ ), aunque este efecto fue solo marginalmente significativo ( $p < 0,10$ ).”

Después de determinar que hay un efecto principal de *grupo* y encontrar una interacción significativa, se recurre a los efectos simples para interpretarla. En primer lugar se llega a la conclusión de que “*Los sujetos del grupo control recordaron mas palabras (negativas-depresivas, positivas-depresivas y positivas-controles) que los sujetos depresivos.*” Es decir, se hace referencia a un efecto principal como si se estuviese interpretando la interacción. A continuación se comentan los efectos simples de *tipo de palabras* en cada uno de los niveles de *grupo*, lo que tampoco es una manera de interpretar o desglosar la interacción.

**Tabla 4.7. Tomada del original (*Psicothema*, 2001)**

<p><i>Tabla 2</i> Medias y desviaciones típicas del número de palabras recordadas correctamente en función del factor Grupo (Depresivo vs Control No-depresivo) y Tipo de palabra (negativa-depresiva, positiva-depresiva y positiva-control)</p>				
Tipo palabra	Grupo			
	Depresivos		Control	
	M	D	M	SD
Negativa-Depresiva	1.33	1.18	3.00	1.69
Positiva-Depresiva	0.27	0.80	5.87	1.96
Positiva-Control	0.33	0.82	4.13	2.59

Una forma correcta de interpretar la interacción global es que la diferencia que existe entre el numero de palabras recordadas por el grupo control y el grupo depresión difiere entre al menos dos niveles de *tipo de palabra* (independientemente de que la diferencia entre el grupo control y el grupo depresión sea significativa en todos los niveles de *tipo de palabras*). Respecto al desglose de la interacción se puede avanzar que la diferencia entre el número de palabras positivas-depresivas recordadas entre el grupo control y el grupo depresión es mayor que la diferencia entre el número de palabras negativas-depresivas recordadas entre el grupo control y el grupo depresión. Otros posibles aspectos de la interacción requieren un análisis formal para determinar si existe o no significación estadística.

**Ejemplo 24)** (*Psicothema*, 2004) Los autores estudian la velocidad en la tarea de nombrar en un grupo de niños con dificultades en el aprendizaje de la lectura, comparándolos con dos grupos control, uno de la misma edad y un grupo de menor edad pero igualado en capacidad lectora. Se trata de un diseño de 2 factores: uno intrasujetos, *tipo de tarea* (con cuatro niveles: colores; números, dibujos; letras) y otro intersujetos, *grupo* (con tres niveles: dificultad en el aprendizaje de la lectura [DAL4°];



control de la misma edad [BL4°]; control de la misma capacidad lectora [BL2°]). Los resultados se presentan en la Figura 1 del artículo que se reproduce en la Figura 4.20.

Los investigadores encontraron efectos principales significativos de *grupo* y *tarea*, y a continuación señalan (pág. 444):

*“No obstante estos efectos estaban mediatizados por una interacción grupo  $\times$  tipo de tarea  $F(6, 180) = 0,343$ ;  $p < 0,05$ ;  $\eta^2 = 0,103$ . Los análisis a posteriori de los efectos simples mostraron que existen diferencias significativas entre los alumnos con DAL y los BL de 4° en los tiempos de ejecución en todas las tareas: nombrar colores  $F(1, 92) = 25,5$ ,  $p < 0,001$ ; nombrar números,  $F(1, 92) = 27,29$ ,  $p < 0,001$ ; nombrar dibujos,  $F(1, 92) = 29,41$ ,  $p < 0,001$ , y; nombrar letras,  $F(1, 92) = 20,08$ ,  $p < 0,001$ . De igual manera, los BL de 4° fueron significativamente más rápidos que los BL de 2° en todas las tareas: colores  $F(1, 92) = 17,10$ ,  $p < 0,001$ ; números,  $F(1, 92) = 22,99$ ,  $p < 0,001$ ; dibujos,  $F(1, 92) = 11,93$ ,  $p < 0,001$ , y letras,  $F(1, 92) = 23,91$ ,  $p < 0,001$ . Sin embargo, no encontramos diferencias significativas en ninguna de las tareas entre los sujetos con DAL de 4° nivel y los lectores de 2° nivel ( $F > 1$ ). En definitiva, los BL de 4° nivel fueron significativamente más rápidos que los dos grupos restantes en las cuatro tareas de nombrar (ver Figura 1).*

En el párrafo anterior los autores tratan de desglosar la interacción recurriendo a los efectos simples. Sin embargo, como estos no arrojan luz sobre cómo se produce la interacción se concluye el efecto principal identificado anteriormente. En el caso de que la interacción sea significativa (lo que no resulta claro pues o hay un error en el estadístico  $F$  o un error en el valor  $p$ ), un desglose adecuado posiblemente encontraría diferencias significativas en números y dibujos entre DAL4° y BL4°.

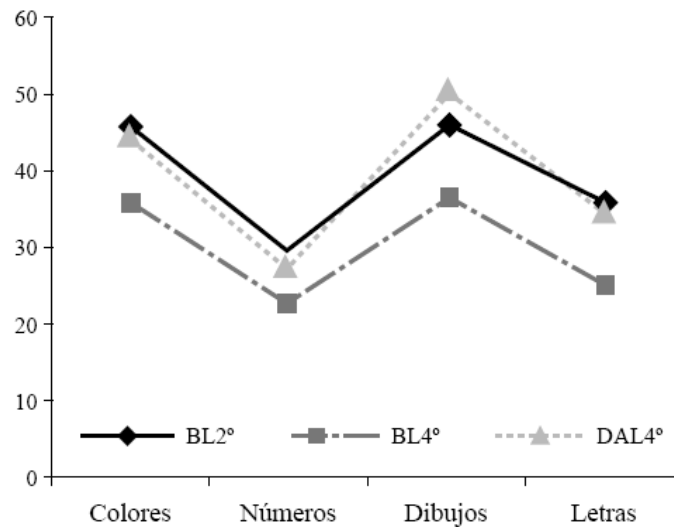
Figura 4.20. Tomada del original (*Psicothema*, 2004)


Figura 1. Tiempos de respuestas (en sg) de los tres grupos en cada una de las tareas de nombrar

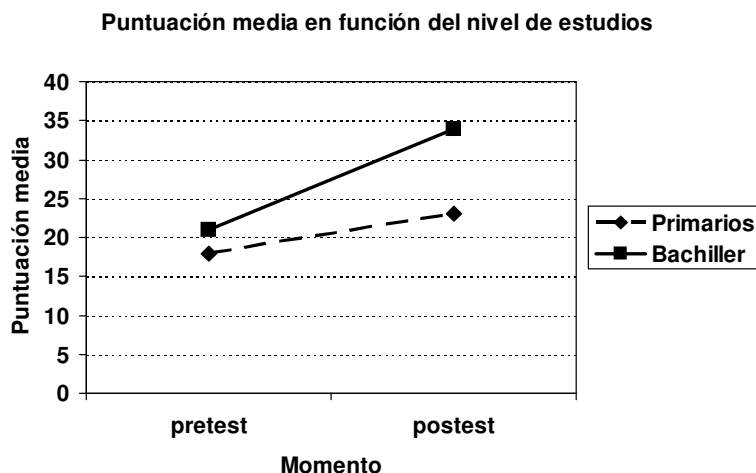
**Ejemplo 25)** (*Psicothema*, 2000) Los autores estudiaron la mejora en las puntuaciones mediante aprendizaje en función del nivel de estudios y la edad. Se trata de un diseño de tres factores: dos factores intersujetos, *nivel de estudios* (con dos niveles: estudios primarios; bachiller) y *grupo de edad* (con tres niveles: <40 a.; 40 – 49 a.; ≥ 50 a.), y un factor intrasujetos, *momento* (con dos niveles: pretest; posttest). Se estudiaron por separado las interacciones *momento* × *nivel de estudios* y *momento* × *grupo de edad*. El modelo de tres factores no ha sido objeto de estudio por los autores y no existe ningún comentario acerca de la posible interacción entre los tres factores.

En el apartado resultados (pág. 366) se señala respecto a la interacción momento por nivel de estudios:

“... Los resultados obtenidos en dicho análisis de varianza fueron: 1) En la variable intersujetos se obtuvo un valor de  $F(1, 109) = 24,48$ ,  $p < 0,001$ . 2) En la variable intrasujetos  $F(1, 109) = 24,48$ ,  $p < 0,001$ . Y 3) en la interacción  $F(1, 109) = 24,48$ ,  $p < 0,001$ . A la vista de los resultados se observa que las diferencias entre los niveles de estudio resultan estadísticamente significativas.” (cursiva mía).

En el párrafo anterior, o bien no se comenta la interacción o bien se comenta como efecto principal. En la Figura 4.21 se presenta un extracto de los resultados relacionados con la interacción *momento*  $\times$  *nivel de estudios*.

**Figura 4.21.** Preparada a partir de la Tabla 3 del original (*Psicothema*, 2000)



Respecto a la interacción *momento*  $\times$  *grupo de edad*, los resultados se presentan en la Tabla 4 del artículo referenciado, que se reproduce en la Tabla 4.8. No se efectúa, o al menos no se comenta, un análisis de la interacción y tan sólo se informa de las pruebas *post hoc* (pág. 366).

*“Al comparar los grupos mediante la prueba de Scheffé se obtuvieron diferencias significativas en el posttest entre los grupos de las personas menores de 40 años y las personas de edad comprendida entre 40 y 49 años,  $p < 0,001$ ; también se encontraron diferencias entre las personas menores de cuarenta años con las personas mayores de cincuenta años,  $p < 0,001$ .” (cursiva mía).*

En este caso no se informa de si la interacción es o no significativa y tan sólo se informa de los efectos simples en uno de los niveles del segundo factor. Por tanto, no se está prestando atención al efecto de la interacción.

Tabla 4.8. Tomada del original (*Psicothema*, 2000)

<p>Tabla 4</p> <p>Puntuaciones medias y desviación típica de los grupos de edades</p>						
Pretest			Posttest			Δ Media
	N	Media	D.T.	Media	D.T.	
< 40	37	21,86	7,79	31,35	7,18	9,49
40-49	57	17,56	6,60	24,12	8,90	6,56
≥50	19	17,57	6,56	22,57	7,47	5,00

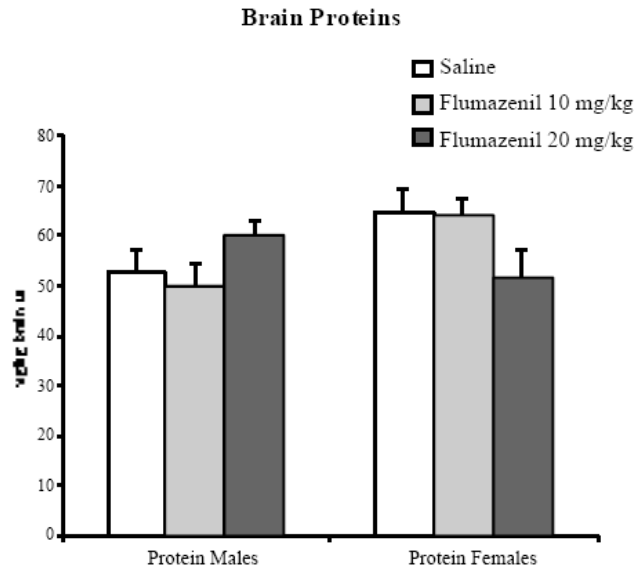
**Ejemplo 26)** (*Psicothema*, 2000) Los autores estudiaron los efectos neuroquímicos sobre el desarrollo de la exposición postnatal a flumazenil en ratones machos y hembras. Se trata de un diseño con dos factores intersujetos: *sexo* (con dos niveles: machos; hembras) y *tratamiento* (con tres niveles: solución salina; flumazenil 10 mg/kg; flumazenil 20 mg/kg). Los resultados se presentan en la Figura 1 del artículo que se reproduce en la Figura 4.22. En el apartado resultados (pág. 623) se señala:

“Only the interaction between Sex × Treatment [ $F(2, 144) = 4.952$ ;  $p < 0.0083$ ] was significant with respect to brain proteins. *Simple main effects show that Sex was significantly different in the saline and flumazenil 10 mg/kg groups ( $p < 0.032$  and  $p < 0.017$ , respectively). The effect of treatment was only evident in females ( $p < 0.025$ ). An ANOVA in the female treatment groups was performed [ $F(2, 72) = 7.686$ ;  $p < 0.0410$ ] but Newman-Keuls analysis did not show differences between groups* (cursiva mía).

La descripción de los efectos simples no ayuda en absoluto a entender la interacción, sino que confunde a los investigadores. El hecho de que existan diferencias entre los niveles medios de proteínas de machos y hembras en los grupos solución salina y flumazenil 10 mg/kg es irrelevante para explicar la interacción. El hallazgo relevante es que la diferencia que se produce entre 10 y 20 mg/kg de flumazemil en machos es de

sentido contrario que la diferencia que se observa entre 10 y 20 mg/kg de flumazenil en hembras, siendo esta diferencia la mayor de todas.

**Figura 4.22. Tomada del original (*Psicothema*, 2000)**



*Figure 1. Means ( $\pm$ S.E.) of brain proteins of male and female pups of the first five days of postnatal life after a single treatment on the day of birth with control or one of two doses of flumazenil*

Como por otra parte no parece haber diferencias entre solución salina y flumazenil 10 mg/kg en ningún sexo, si un análisis desglosado de la interacción ratificase esta impresión, habría que concluir que mientras flumazenil 10 mg/kg parece no tener influencia sobre el nivel de proteínas, flumazenil 20 mg/kg actúa de manera diferente sobre el nivel de proteínas en función del sexo de los ratones, que tiende a aumentar en machos mientras que tiende a disminuir en hembras.

**Ejemplo 27** (*Psicothema*, 2004) Los autores estudiaron la relación entre la memoria de trabajo (MT) y los procesos de inhibición en la memoria episódica, utilizando una tarea de interferencia proactiva (IP). Se trata de un diseño con 4 factores: 1 factor intersujetos, *amplitud de MT* (con dos niveles: alta amplitud; baja amplitud) y 3 factores intrasujetos, *tipo de ensayo*, (con dos niveles: alta interferencia; baja interferencia), *carga de*

*memoria* (con dos niveles: baja carga [2 palabras]; alta carga [4 palabras]) y *tipo de respuesta* (con dos niveles: sí; no). Las variables dependientes eran el TR y número de errores.

Los resultados del experimento 2 presentado por los autores se muestran en la Tabla 2 del artículo, la cual se reproduce en la Tabla 4.9. En relación con los tiempos de respuesta en la pág. 191 se dice:

“...No se observó interacción entre la variable tipo de respuesta y ninguna de las otras variables (todas las  $p > .05$ ). Debido a esta ausencia de interacción, se colapsaron los datos y se llevo a cabo un ANOVA 2 (amplitud de MT)  $\times$  2 (tipo de ensayo)  $\times$  2 (carga de memoria)

Este análisis reveló un efecto significativo del tipo de ensayo.....Por otro lado, *ninguna de las interacciones fue significativa  $p > .05$ . A pesar de que la interacción Tipo de ensayo  $\times$  Carga de memoria  $\times$  Amplitud no alcanzó significatividad estadística, realizamos un análisis por separado para cada uno de los grupos de amplitud, ya que existían hipótesis a priori sobre posibles diferencias entre los grupos que justificaban este análisis.*

Los resultados del análisis *en el grupo de Baja amplitud mostraron una diferencia significativa entre ensayos AI y ensayos BI,  $F(1, 10)=12,59$ ,  $MS=13412$ ,  $p > 0,01$ , los tiempos de respuesta en los ensayos AI fueron 124 ms más lentos que en los ensayos BI, lo que refleja el efecto de la interferencia.....*

Por el contrario, *los resultados del grupo de Alta amplitud no mostraron interferencia proactiva. Es decir, las diferencias en los tiempos de respuesta entre los ensayos AI y BI no fueron significativas  $F(1, 10)=3.43$ ,  $MS=15650$ ,  $p > .05$ . ...*

En resumen, si consideramos los resultados obtenidos en ambos grupos podemos concluir que, aunque todos los participantes se vieron afectados por la carga de la memoria, *solamente los participantes de baja amplitud de MT mostraron efectos de interferencia proactiva, mientras que los pacientes de alta amplitud de MT no experimentaron interferencia en memoria episódica.*” (cursiva mía).

En el primer párrafo se dice que se agrupan los niveles de uno de los factores, por lo que el diseño queda reducido a  $2 \times 2 \times 2$ . En el segundo, que no existen interacciones

significativas entre los tres factores, ya sea considerados conjuntamente o de dos en dos, pero aun así se van a estudiar independientemente los dos grupos de amplitud de memoria puesto que existen hipótesis “a priori”. En el tercero, que se ha encontrado un efecto significativo de la carga de memoria en el grupo de baja amplitud. En el cuarto, que no se ha encontrado un efecto significativo de la carga de memoria en el grupo de alta amplitud. Y, en el quinto, que existe una interacción que ya se había señalado como no significativa en el párrafo segundo. Si se tiene en cuenta que la hipótesis formulada por los investigadores para el experimento 2 es “*que los participantes de alta amplitud de MT mostrarían una menor susceptibilidad a la interferencia proactiva que los participantes de baja amplitud de MT*” (pág. 190), es decir, que existe interacción entre amplitud de MT e interferencia proactiva, este ejemplo pone en evidencia una vez más que el análisis de los efectos simples puede llevar a conclusiones diferentes que el análisis de interacción.

**Tabla 4.9. Tomada del original (*Psicothema*, 2004)**

<p><i>Tabla 2</i>  Tiempos de Respuesta promedio (TR, en milisegundos), porcentaje de errores (% Errores) y Desviaciones típicas (entre paréntesis) en la tarea de interferencia episódica (Experimento 2), en función de la Interferencia (Baja vs. Alta), la Amplitud de MT de los participantes (Baja y Alta) y la Carga (Baja vs. Alta)</p>				
	Baja interferencia		Alta interferencia	
	TR	% Errores	TR	% Errores
Baja amplitud				
Baja carga	1496.32 (309.19)	15.45 (29.11)	1607.31 (351.68)	8.18 (17.79)
Alta carga	1629.68 (339.43)	20.00 (26.83)	1766.45 (348.91)	24.54 (26.59)
Alta amplitud				
Baja carga	1483.56 (323.97)	4.54 (6.87)	1565.46 (331.98)	5.45 (6.87)
Alta carga	1587.91 (338.23)	10.00 (10.95)	1645.68 (340.38)	17.27 (14.20)

### C) Ejemplos de interpretación correcta de la interacción de primer orden

La interacción entre dos factores se interpreta correctamente cuando se entiende como comparación entre diferencias o cuando se argumenta que la combinación de efectos difiere de la suma de efectos. La interpretación correcta de la interacción tiene diferentes implicaciones en diseños  $2 \times 2$  y en diseños en que alguno de los factores tiene más de dos niveles. En el primer caso basta con hacer una interpretación adecuada de la  $F$  de la interacción a la vista de una tabla de medias o un gráfico de líneas, ya que la comparación entre diferencias involucra a las cuatro celdas disponibles, mientras que en el segundo caso es necesario realizar análisis adicionales para determinar entre qué niveles de ambas variables independientes se producen las diferencias entre diferencias que generan la interacción detectada.

En los artículos revisados se han encontrado ejemplos de diseños  $2 \times 2$  en los que la interacción se analiza e interpreta correctamente. En estos ejemplos la interacción se interpreta bien como diferencia entre diferencias, bien como una evolución diferencial entre dos grupos sometidos a diferentes tratamientos en los estudios pre-post, bien como un efecto combinado diferente de la suma de los de sus componentes. Sin embargo, no se ha encontrado ningún artículo en que, utilizando diseños de 2 factores con más de dos niveles en alguno de ellos, se hayan realizado análisis adicionales adecuados para determinar qué diferencias entre diferencias estaban implicadas en la interacción global encontrada ni, por tanto, se haya realizado una interpretación adecuada.

**Ejemplo 28)** (*Psicothema*, 2002) Los autores estudiaron la influencia de las interrupciones en el recuerdo del desarrollo de un guión. Se trata de un diseño con dos

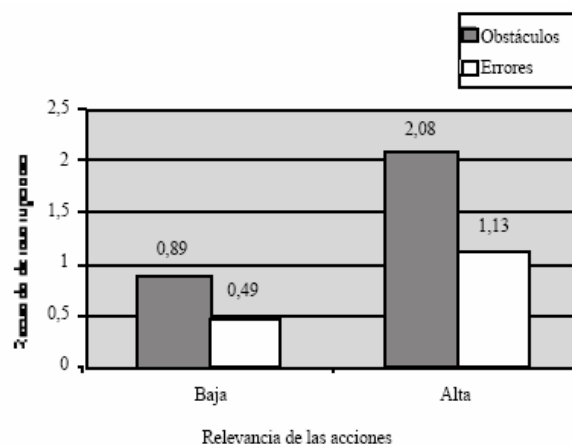


factores intrasujetos: *tipo de interrupción*, (con dos niveles: obstáculos; errores) y *relevancia* [de la acción dentro del guión] (con dos niveles: alta; baja). Los resultados se presentan en la Figura 1 del artículo que se reproduce en la Figura 4.23. En el apartado resultados (pág. 779) se señala:

[En primer lugar se comentan los efectos principales]... El efecto de estas variables fue modulado por una interacción menor entre ambas  $F(1, 98) = 5,40$ ,  $p = 0,022$ ,  $MSe = 1,41$ ,  $\eta = 0,05$ . ...*La interacción muestra que la diferencia entre los niveles de recuerdo de obstáculos y errores es mayor cuando están relacionados con las acciones de alta relevancia* (obstáculos  $M = 2,08$ ;  $DT = 1,33$  y errores  $M = 1,13$ ;  $DT = 1,20$ ) *que cuando están relacionados con los de baja relevancia* (obstáculos  $M = 0,89$ ;  $DT = 1,02$  y errores  $M = 0,49$ ;  $DT = 0,77$ ) (cursiva mía).

En este caso la interacción se está interpretando correctamente como diferencia entre diferencias. Se está diciendo que la diferencia entre recuerdo de obstáculos y errores es mayor en acciones de alta relevancia que en acciones de baja relevancia. Por tratarse de un diseño  $2 \times 2$ , esta interpretación no necesita ser acompañada de ningún contraste adicional.

**Figura 4.23. Tomada del original (*Psicothema*, 2002)**



**Figura 1.** El recuerdo de las interrupciones (obstáculos y errores) en función de la relevancia de las acciones (alta y baja) donde se produce la interrupción

**Ejemplo 29** (*Psicothema*, 2002) Los autores estudiaron la influencia de un programa de desarrollo de la creatividad sobre las subpruebas del Test de Pensamiento Creativo de Torrance (TTCT) en un contexto escolar. Se trata de un diseño con dos factores: uno intrasujeto, *momento* (con dos niveles: pretest; postest) y otro intersujetos, *grupo* (con dos niveles: control; experimental). Los resultados se presentan en la Tabla 1 del artículo referenciado, que se presenta en la Tabla 4.10. En el apartado resultados (pág. 412) se comenta:

“Como podemos observar, no se producen diferencias entre los grupos experimental y control en ninguno de los aspectos de la creatividad, como evidencia la falta de significatividad de la prueba  $F$  entre sujetos. Por el contrario, se observa un efecto claramente significativo de la variable intrasujetos en cada variable. Esto es, la puntuación de ambos grupos tomada en conjunto cambia significativamente a través desde el inicio al final del programa, tanto en fluidez ( $F_{(77, 1)} = 11,62, p = .001$ ), como en la flexibilidad ( $F_{(77, 1)} = 17,36, p = .001$ ), la originalidad ( $F_{(77, 1)} = 19,15, p = .001$ ), la elaboración ( $F_{(77, 1)} = 44,98, p = .001$ ) y el total de la prueba ( $F_{(77, 1)} = 33,79, p = .001$ )

Para conocer si este cambio se produce por igual en ambos grupos, el experimental y el control, o en alguno de ellos de forma diferente al otro, observamos el efecto de la interacción. El efecto de la interacción grupo  $\times$  momento de evaluación resulta significativo en algunos casos, *indicando que el cambio no se produce del mismo modo en el grupo experimental y en el grupo control*. La interacción resulta significativa en el caso de las variables de flexibilidad ( $F_{(77,1)} = 4,01, p = 0,043$ ), originalidad ( $F_{(77,1)} = 4,33, p = 0,039$ ) y en la puntuación total ( $F_{(77,1)} = 4,10, p = 0,041$ ). En todas las variables de creatividad, pero especialmente en aquellas en las que la interacción es significativa, se observa un cambio en el grupo experimental mayor que en el grupo control, desde el inicio al final del programa de desarrollo de la creatividad.”

En este párrafo la interacción se interpreta correctamente como diferencia entre la evolución de los grupos experimental y control a lo largo del estudio. Después de describir los efectos principales se pasa a describir los efectos de la interacción entre

variables, destacando que el cambio en el grupo experimental fue mayor que en el grupo control en todas las variables de creatividad y que esta diferencia en la evolución de ambos grupos fue mayor para aquellas variables dependientes en que se encontró una interacción significativa.

Es de destacar que en este estudio no se hizo ningún contraste adicional a la prueba  $F$  y que los efectos simples ni siquiera se mencionan, no siendo necesarios para interpretar adecuadamente la interacción.

**Tabla 4.10. Tomada del original (*Psicothema*, 2002)**

<p><i>Tabla 1</i> Resultados del análisis factorial de la varianza realizado sobre las puntuaciones de cada uno de los aspectos de creatividad en los grupos de Educación Infantil</p>						
Efecto	Entre sujetos		Intra sujetos		Interacción	
Variables	F	P	F	P	F	P
Fluidez	.03	.86	11.62	.001	1.21	.274
Flexibilidad	.01	.94	17.36	.001	4.01	.043*
Originalidad	.02	.88	19.15	.001	4.33	.039*
Elaboración	1.91	.17	44.98	.001	.35	.558
Total	.11	.73	33.79	.001	4.10	.042*
* = significación estadística						

***Ejemplo 30)*** (*Psicothema*, 2004) Los autores estudiaron la influencia de la intervención en procesos de revisión del mensaje sobre procesos implicados en la escritura. Se trata de un diseño con dos factores, uno intrasujetos, *momento* (con dos niveles: pretest; posttest) y otro intersujeto, *grupo* (con dos niveles: control, experimental). En el apartado resultados (pág.198) se comenta:

*“Serán relevantes para este diseño, los cambios que tengan lugar en el transcurso del tiempo, solo si se confirma que los cambios son significativamente mayores en el grupo experimental, es decir, cuando existe interacción entre el tratamiento y el tiempo.”*  
(cursiva mía).

La Tabla III del artículo, que recoge las medias de los grupos control y experimental en el pre-test y post-test, la  $F$  de la interacción y su significación, y el tamaño del efecto, se reproduce en la Tabla 4.11. El aspecto a destacar en esta tabla es que no se analizan los efectos simples, ya que en un diseño  $2 \times 2$ , no es necesario ningún contraste adicional para explicar una interacción significativa.

**Tabla 4.11. Tomada del original (*Psicothema*, 2004)**

Tabla III Mejoras estadísticamente significativas en productividad y en coherencia en la tarea de redacción, y en las tareas de reflexividad hacia la escritura. Para este diseño, lo relevante son los cambios del pre al post pero sólo si se da en mayor medida en el experimental que en el control (interacción)									
Variable	Intervención			Control			Pre-post x tratamiento		
	N	Media pre (Mia)	Media post (Mip)	N	Media pre (Mca)	Media post (Mcp)	$F_{(1, 90)}$	p	$\eta^2$
<b>Parámetros de productividad (redacción)</b>									
Redacción: Palabras de contenidos	16	39.37	54.56	76	38.63	35.92	8.66	0.0041	0.0878
Redacción: Palabras funcionales	16	19.93	28.56	76	23.31	20.35	9.09	0.0033	0.0917
Redacción: Determinantes	16	10.75	19.06	76	14.44	14.56	7.07	0.0092	0.0728
Redacción: Número de palabras total	16	70.06	101.56	76	77.69	74	5.38	0.0225	0.0564
<b>Indicadores de coherencia (redacción)</b>									
Redacción: Coherencia: Enlaces	16	0.18	0.31	76	0.44	0.28	3.68	0.0581	0.0005
Redacción: Coherencia: Párrafos	16	0.18	0.68	76	0.40	0.34	10.38	0.0017	0.0278
Redacción: Coherencia: Hilo Argumental	16	0.43	0.93	76	0.82	0.72	14.94	0.0002	0.0734
Redacción: Coherencia: Total	16	1.75	2.87	76	2.59	2.26	14.74	0.0002	0.0349
<b>Indicadores de reflexividad</b>									
Estilos Cognitivos 1: Aciertos	16	9.18	26.81	76	19.94	25.43	29.82	0.0000	0.2489
Estilos Cognitivos 1: Errores	16	3.06	1.31	76	1.85	2.43	4.97	0.0282	0.0524
Estilos Cognitivos 2: Sistema de búsqueda	16	0.37	0.81	76	0.59	0.72	3.41	0.0679	0.0365
Mia= Media del grupo de intervención antes; Mip= Media del grupo de intervención post; Mca= Media del grupo de control antes; MCP= Media del grupo de control post.									

El análisis de los efectos simples podrá tener o no interés en función de los objetivos del estudio, pero no explica la interacción. En este sentido los autores del estudio declaran acertadamente estar interesados en los cambios que se producen a lo largo del tiempo “solo si se confirma que los cambios son significativamente mayores en el grupo experimental, es decir, cuando existe interacción entre el tratamiento y el tiempo.”

Un caso particular en que, habiendo o no interacción, el análisis de los efectos simples tiene sentido es cuando nos encontramos ante comparaciones planeadas guiadas por la hipótesis de investigación. A continuación se presentan dos ejemplos.

**Ejemplo 31** (*Psicothema*, 2003) Los autores utilizaron una versión modificada de los experimentos de olvido dirigido para estudiar si el olvido dirigido afecta a las falsas memorias. En el paradigma de olvido dirigido los sujetos son expuestos a dos listas de nombres y mientras a un grupo se le pide que olvide la lista 1 para concentrarse en la lista 2 al otro se le pide que recuerde las palabras de ambas listas. Se introduce un nuevo factor consistente en si las palabras de la lista 1 están relacionadas o no con una palabra crítica no explícitamente presente en la lista. Se trata de un diseño de tres factores  $2 \times 2 \times 2$ , con dos factores intersujetos, *grupo* (con dos niveles: olvida [OLV]; recuerda [REC]) y *condición* (con dos niveles: relacionada [REL]; no relacionada [NR]), y un factor intrasujetos, *lista* (con dos niveles: lista 1 [L1]; lista 2 [L2]). En la Tabla 1 del estudio, que se reproduce en la Tabla 4.12, se presentan parte de los resultados del estudio. En el apartado resultados (pág. 8) se indica:

“A continuación analizamos la variable dependiente proporción de aciertos en la condición no relacionada. Para ello llevamos a cabo un ANOVA mixto 2 (OLV vs REC)  $\times$  2 (L1 vs L2) siendo la primera variable intersujetos y la última intrasujetos. Las proporciones medias de aciertos se presentan en la parte izquierda de la Tabla 1. Dicho ANOVA determinó como significativa la interacción de ambas variables  $F(81, 50) = 9,233$ ,  $M_{ce} = 3,037$ ,  $p < 0,01$ . *A continuación analizamos mediante pruebas t los dos patrones de diferencias que evidencian un efecto de olvido dirigido, independientemente de si dicha interacción es o no significativa (ver conway et al., 2000), tal como suele hacerse con las comparaciones planificadas a priori (ver Keppel, 1991): a) en el grupo olvida es de esperar que su recuerdo sea significativamente peor en la lista 1 que en la lista 2; b) en el grupo olvida es de esperar que su recuerdo de la lista 1 sea peor que el del grupo recuerda sobre la misma lista (cursiva mía).*

(Este mismo abordaje se repite en el grupo de condición relacionada).

El aspecto a destacar de este planteamiento es que la comparación de dos celdas puede estar justificada cuando está avalada por hipótesis *a priori* que requieren la planificación

previa de dicha comparación, mientras que no está justificada cuando se utilizan como explicación de una interacción.

**Tabla 4.12. Tomada del original (Psicothema, 2003)**

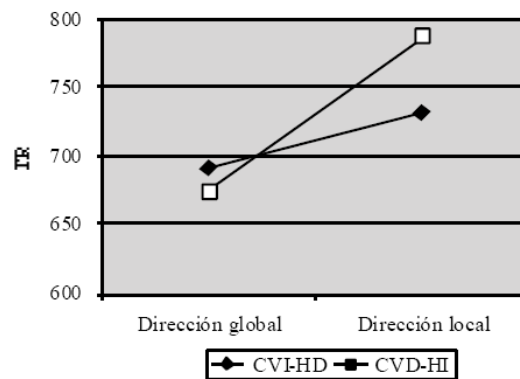
<p><i>Tabla 1</i> Proporciones medias de <i>aciertos</i> por condición en la tarea de recuerdo del experimento 1</p>				
PALABRAS:	NO RELACIONADAS		RELACIONADAS	
	Olvida	Recuerda	Olvida	Recuerda
Lista 1	0.46	0.53	0.62	0.60
Lista 2	0.52	0.42	0.52	0.39

**Ejemplo 32) (Psicothema, 2000)** Los autores estudiaron la asimetría hemisférica en el procesamiento de la información de los rasgos globales y locales de los estímulos visuales. Algunos experimentos previos sugieren que el hemisferio derecho (HD) es superior en el análisis de los rasgos globales y el hemisferio izquierdo (HI) es superior en el análisis de los rasgos locales, mientras que algunos autores no encuentran dicha asimetría. Otros estudios dan apoyo empírico a dicha dicotomía sólo bajo determinadas condiciones de tiempo de exposición estimular y con estímulos de determinado tamaño y número de elementos locales. Los autores pretenden determinar si la asimetría es extensible a tareas de atención selectiva. Para ello realizaron un experimento de atención dividida en el que los estímulos se presentaban en el campo visual derecho (CVD-HI) o izquierdo (CVI-HD) y el sujeto debía indicar si determinada figura (cuadro abierto a la izquierda) estaba presente en el estímulo al nivel que solicitasen las instrucciones: nivel global (figura completa) o nivel local (cada uno de los componentes de la figura completa). Se trata de un diseño de tres factores intrasujetos: *dirección de la atención* (con dos niveles: global; local), *campo visual* (con dos niveles: CVD-H; CVI-HD) y *congruencia de las señales global y local* (con dos niveles: congruente;

incongruente). Los resultados se muestran en la Figura 2 del artículo referenciado, que se reproduce en la Figura 4.24. En el apartado resultados (pág. 16) se señala:

“.... Con respecto a las interacciones, sólo ha resultado significativa la interacción entre la dirección de la atención y el campo visual en TR [ $F(1, 18) = 5,22$ ,  $MSE = 49147,97$ ;  $p = 0,035$ ]. Los resultados se muestran en la Figura 2. Con respecto al CVI-HD, no se encontraron diferencias en la respuesta a los niveles globales y locales ( $t(1, 18) = 1,26$ ;  $p > 0,05$ ). Sin embargo, el HI es más rápido en procesar los rasgos globales ( $t(1, 18) = 3,57$ ;  $p < 0,05$ ). Por otro lado, la comparación entre ambos indica que cuando la atención se dirige al nivel global, ambos hemisferios son igualmente rápidos ( $t(1, 18) = 0,48$ ;  $p > 0,05$ ), mientras que cuando la atención se dirige a nivel local, el HD es más rápido que el HI ( $t(1, 18) = 1,8$ ;  $p < 0,05$ ).

**Figura 4.24.** Tomada del original (Psicothema, 2000)



**Figura 2.** Media en TR en función de la dirección de la atención y campo visual

Las dos últimas comparaciones señaladas en el párrafo anterior están explícitamente definidas en los objetivos del estudio, por lo que se trata de comparaciones planeadas y, por tanto, de interés independientemente de que la interacción entre los dos factores haya resultado o no significativa.

## 4.2. Ejemplos de errores que se cometen al analizar la interacción de segundo orden

En este apartado se presentan ejemplos de los mismos bloques definidos en la interacción entre dos factores: errores consistentes en no analizar o no interpretar la interacción, errores consistentes en analizar y/o interpretar la interacción recurriendo a los efectos simples y ejemplos en que la interacción entre tres factores está correctamente analizada e interpretada.

### A) Ejemplos de errores consistentes en no analizar y/o no interpretar la interacción de segundo orden

**Ejemplo 33) (Psicothema, 2005)** Los autores estudiaron cómo afecta la inclusividad al favoritismo endogrupal y la deshumanización del exogrupo, midiendo las actitudes positivas y negativas hacia el endogrupo y al exogrupo. Para ello utilizaron un diseño de cuatro factores: con un factor intersujetos, *identidad activada* (con tres niveles: supranacional; nacional; regional) y tres factores intrasujetos, *grupo estímulo* (con dos niveles: endogrupo; exogrupo), *tipo de emoción* (con dos niveles: sentimientos; emociones) y *valencia* (con dos niveles: positiva; negativa). Para estudiar la infrahumanización del exogrupo utilizaron un ANOVA de tres factores. El apartado resultados (pág. 444) señala:

“La infrahumanización del exogrupo. En primer lugar, se llevo a cabo un ANOVA de 2 (tipo de emoción: sentimiento vs. emoción)  $\times$  2 (valencia: positiva vs. negativa)  $\times$  2 (grupo estímulo: endogrupo vs. exogrupo). Los resultados confirman los datos de



Leyens y cols. (2000), ya que *se produjo una interacción significativa de las tres variables* ( $F_{(1, 67)} = 14,9, p < 0.001$ ). *Esta interacción matizó los efectos principales obtenidos y las interacciones dobles de las variables grupo estímulo × tipo de emoción* ( $F_{(1, 67)} = 22,3, p < 0.001$ ) *y grupo estímulo × valencia de la emoción* ( $F_{(1, 67)} = 14,7, p < 0.001$ )” [a continuación se realiza una descripción de las interacciones dobles]

En este párrafo no se explica en qué consiste la interacción entre las tres variables y posteriormente este hallazgo no se vuelve a mencionar en el apartado de resultados.

**Ejemplo 34)** (*Psicothema*, 2005) Los autores estudiaron la equivalencia entre las estimaciones de tiempo de llegada de un vehículo en función de la velocidad de éste, de la distancia, de que el desarrollo del test se realizase en un automóvil o en imágenes de video, y del sexo del participante. Se trata de un diseño de cuatro factores: un factor intersujetos, *sexo* (con dos niveles: hombre; mujer) y tres factores intrasujeto: *test* (con dos niveles: vehículo real; video), *velocidad* (con cuatro niveles: 60; 80; 100; 120 km/h) y *distancia* (con cuatro niveles: 75; 100; 125; 150 m). En el apartado resultados (pág. 115) se indica que no se encontró un efecto principal de sexo ni una interacción de sexo con ninguna otra variable, por lo que el análisis subsiguiente se centra en los tres factores con medidas repetidas.

“There was a significant effect of speed  $F(3, 114) = 6.349, p < .001$ , indicating that the higher speed, the lower error, going from -1,27 at 60 kph to -,88 at 120 kph. There was no interaction between test and speed,  $F(3, 114) = .740, p > .05$ . The effect of distance was not significant  $F(3, 114) = 1.015, p > .05$  nor was its interaction with test,  $F(3, 114) = 1.367, p > .05$ . There was a slight interaction between distance and speed  $F(9, 342) = 2.115, p < .05$ , which appeared in both tests, although the interaction Distance × Speed × Test was not significant  $F(9, 342) = 1.138, p > .05$ .

Upon analysing the errors for each test individually, *the effect of speed was significant at each one: for the vehicle*  $F(3, 114) = 5.76, p < .05$ , *partial*  $\eta^2 = .132$ ;

*for the video,  $F(3, 114) = 3.99, p < .05$ , partial  $\eta^2 = .095$ . By itself, distance did not reveal significant effects in any test considered individually: for the vehicle,  $F(3, 114) = 2.01, p > .05$ ; for the video,  $F(3, 114) = 0.90, p = .443$ . (cursiva mía).*

La frase “*there was a slight interaction between distance and speed, which appeared in both tests, **although** the interaction Distance  $\times$  Speed  $\times$  Test was not significant*” implica una interpretación inadecuada. La interpretación correcta sería: como no se produce interacción entre los tres factores, la interacción entre distancia y velocidad se produce de forma similar en los dos niveles de test.

En el segundo párrafo no se toman en consideración los hallazgos del análisis de la interacción. Aunque previamente se ha indicado que no hay interacción significativa entre velocidad y test, ni entre distancia y test, ni entre los tres factores, se analizan los errores para cada test individualmente, encontrándose, como no podía ser de otra manera, los mismos resultados que en análisis de los efectos principales de velocidad y distancia. Estos análisis resultan totalmente innecesarios si los resultados del análisis de la interacción se interpretan adecuadamente.

## **B) Ejemplos de errores en los que la interacción de segundo orden se interpreta recurriendo al análisis de los efectos simples**

Al igual que sucede en el análisis de la interacción entre dos factores, el recurso de utilizar al análisis de los efectos simples para interpretar una interacción significativa también es utilizado en interacciones de orden mayor.

Cuando se considera una interacción significativa de segundo orden, la interpretación correcta consiste en considerar que la interacción entre los dos primeros factores difiere significativamente entre al menos dos niveles del tercer factor. Por tanto,

existen varias formas de interpretación basadas en otros análisis que, siguiendo a Kirk (1995, pág. 443), seguiremos llamando análisis de los efectos simples, pero que no son útiles para explicar la interacción de segundo orden. Dentro de estos análisis de los efectos simples distinguiremos tres tipos diferentes, según la interacción se interprete recurriendo: al análisis de los *efectos simples simples* (por medio de la comparación entre niveles de un factor en las combinaciones de los otros dos factores), al análisis de los *efectos simples* (agrupando los niveles de una variable y analizando los efectos simples en las dos restantes), o al análisis de los efectos *simples de interacción* (analizando las interacciones de dos factores en los niveles del tercer factor, pero sin compararlas formalmente entre ellas).

### **B.1) Interpretación a partir de los efectos simples simples**

**Ejemplo 35) (Psicothema, 2003)** Los autores estudiaron si la diferencia percibida en valores entre el endogrupo y el exogrupo está relacionada con los prejuicios intergrupales. Para ello aplicaron a una muestra de 145 estudiantes de secundaria la escala de prejuicios de Pettigrew y Meertens que contiene dos subescalas, una de prejuicio manifiesto y otra de prejuicio sutil, y la versión española de la escala de valores Schwartz y Bilsky, cuyos 10 campos motivacionales básicos que guían la vida se aplicaron al endogrupo y al exogrupo. Las variables dependientes fueron las puntuaciones en cada uno de los 10 valores estudiados. Los sujetos se dividieron en tres grupos de acuerdo con el tercil en prejuicio sutil y manifiesto, y sólo se compararon el tercil superior (alto nivel de prejuicio) y el inferior (bajo nivel de prejuicio). Se trata por tanto de un diseño de 4 factores: dos factores intersujetos *prejuicio sutil* (con dos

niveles: alto y bajo) y *prejuicio manifiesto* (con dos niveles: alto y bajo) y dos factores intersujetos, *grupo* (con dos niveles: endogrupo y exogrupo) y *valores* (con 10 niveles).

En la introducción (pág. 177) se indica:

“El objetivo general de la presente investigación consiste en analizar la relación existente entre la diferencia percibida en valores entre el endogrupo payo y el exogrupo gitano y el antagonismo actitudinal (prejuicio) hacia este último. Así, el primer objetivo se centra en determinar si en las relaciones payos-gitanos es la similitud en valores lo que determinaría menores niveles de prejuicio, tal como se derivaría de algunos estudios anteriores (Gómez y Ruiz, 2001; Schwartz y Struch, 1989, Struch y Schwartz, 1989), o esta similitud produciría mayor diferenciación y por tanto aptitudes más prejuiciosas hacia los gitanos como proponía la Teoría de la Identidad Social.”

Como las clasificaciones en *prejuicio sutil* y *prejuicio manifiesto* no son excluyentes, no se puede plantear un ANOVA de 4 factores, por lo que se desarrollan 2 ANOVAs, uno para *prejuicio sutil* y otro para *prejuicio manifiesto*, cada uno con 3 factores, *nivel de prejuicio* (alto, bajo)  $\times$  *grupo* (endogrupo, exogrupo)  $\times$  *valores* (10 niveles). La interacción entre los tres factores resultó significativa cuando como variable independiente se tomó el nivel de prejuicio sutil. Dicha interacción se interpreta sobre la Tabla 1 del artículo que se reproduce en la Tabla 4.13. En las págs. 179-80 se indica:

“El ANOVA  $2 \times 2 \times 10$  similar al anterior, pero tomando como variable entregupos el nivel de prejuicio sutil, apoyó los resultados mostrados con el prejuicio manifiesto...

...Además esta interacción se vio modulada por el Tipo de Valores, como lo muestra la interacción Nivel de prejuicio sutil  $\times$  Grupo de Valoración  $\times$  Tipo de Valores  $F(9, 900) = 3,74; p < .001$ . Con objeto de analizar esta interacción se realizó un ANOVA para cada tipo motivacional en los que se incluyeron como variable entregupos el Nivel de prejuicio sutil (alto vs. bajo) y como variable intraparticipantes el grupo de valoración (endogrupo vs. exogrupo)

Tal como puede observarse en la Tabla 1, los participantes con alto nivel de prejuicio sutil pensaban que el endogrupo daba más importancia que el exogrupo a los tipos de valores «Universalismo», «Benevolencia» y «Conformidad», mientras que los participantes con bajo nivel de prejuicio no percibían diferencias entre el endogrupo y el exogrupo en estos tipos motivacionales. Curiosamente los individuos con alto nivel de prejuicio sutil consideran igualmente motivados al endogrupo y al exogrupo por el valor «Poder», mientras que los de bajo prejuicio consideran menos importante este tipo motivacional para el exogrupo gitano que para el endogrupo payo. Por último, tanto los individuos con alto como con bajo prejuicio sutil percibieron como más importante el «Logro» para el endogrupo que para el exogrupo gitano.”

En los párrafos anteriores se describe: a) que se detecta una interacción significativa entre tres factores; b) que para analizarla se realiza un ANOVA de dos factores en cada uno de los 10 niveles del tercero; c) que la interacción se interpreta a partir de las comparaciones entre pares de celdas. Dichas comparaciones entre pares de celdas del diseño original (sin agrupar niveles) son los *efectos simples simples*.<sup>8</sup> Y ya se ha señalado que los *efectos simples simples* pueden tener interés, pero no ayudan a interpretar una interacción significativa entre tres factores.

Además, en la discusión se deslizan algunos comentarios comparativos entre niveles de prejuicio manifiesto y sutil (págs. 181-2) que no se corresponden con los ANOVAs realizados, ya que en ningún momento se ha realizado una comparación explícita entre prejuicio sutil y manifiesto.

“Sin embargo, cuando nos centramos algo más detenidamente en los valores en que las diferencias percibidas entre payos y gitanos son más marcadas, en función de los niveles de prejuicio manifiesto o sutil que tienen, podemos encontrar matizaciones a la

---

<sup>8</sup> Si en lugar de utilizar comparaciones entre dos celdas para interpretar la interacción entre los tres factores, se hubiesen utilizado las interacciones entre dos factores, también facilitadas en la Tabla 1, sin compararlas formalmente, se trataría de un caso de utilización de los efectos simples de interacción para interpretar la interacción entre tres factores.

supuesta equivalencia entre estos dos tipos de prejuicio en nuestro estudio, lo que quizás aporte algunas pistas acerca de las características que puedan distinguirlos.”

Este párrafo sugiere la existencia de una interacción que no se puede poner a prueba con el diseño utilizado en el estudio, ya que las clasificaciones en prejuicio sutil y prejuicio manifiesto no son exclusivas.

**Tabla 4.13. Tomada del original (*Psicothema*, 2003)**

Tabla 1 Medias y comparaciones de la valoración de los diferentes tipos motivacionales para el endogrupo y el exogrupo en función del nivel de prejuicio sutil					
PREJUICIO SUTIL	Bajo nivel de prejuicio		Alto nivel de prejuicio		
Campo motivacional	Media endogrupo	Media exogrupo	Media endogrupo	Media exogrupo	Interacción Grupo de valoración x Nivel de prejuicio sutil para cada campo de valores
Universalismo ( <i>igualdad, mundo en paz y armonía interna</i> )	4,09	3,84	4,42	3,31 <sup>1</sup>	F(1,100)= 7,44; p<.01
Benevolencia ( <i>ayudar, honesto, no rencoroso, leal, responsable, amistad verdadera, vida espiritual, amor maduro y sentido en la vida</i> )	4,02	4,1	4,27	3,13 <sup>1</sup>	F(1,100)= 17,64; p<.01
Tradicición ( <i>dévoto, acepto mi parte en la vida, humilde, moderado, respeto por la tradición y distanciamiento</i> )	3,51	3,68	3,57	3,67	n.s.
Conformidad ( <i>buenos modales, honrar a padres y mayores, obediente y autodisciplinado</i> )	3,77	3,71	4,31	3,20 <sup>1</sup>	F(1,100)= 10,25; p<.01
Seguridad ( <i>seguridad, orden social, limpieza</i> )	4,02	3,78	4,23	3,54 <sup>1</sup>	n.s.
Poder ( <i>poder social, autoridad y riqueza</i> )	3,88	3,28 <sup>1</sup>	3,96	3,87	F(1,100)= 3,60; p=. .06
Logro ( <i>exitoso, capaz, ambicioso</i> )	4,23	3,77 <sup>1</sup>	4,21	3,27 <sup>1</sup>	n.s.
Hedonismo ( <i>placer, disfrutar de la vida</i> )	4,37	4,24	4,87	4,68	n.s.
Estimulación ( <i>atrevido, vida variada, vida excitante</i> )	3,56	3,56	4,09	4,01	n.s.
Autodirección ( <i>Creatividad, curioso, libertad</i> )	4,11	3,82	4,12	3,71	n.s.

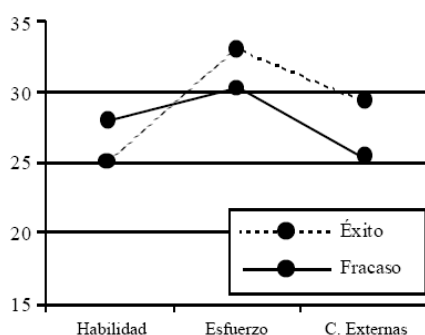
<sup>1</sup> Comparaciones endogrupo-exogrupo significativas al nivel 0,01

**Ejemplo 36) (*Psicothema*, 2000)** Los autores estudiaron las atribuciones causales a distintos resultados en alumnos con y sin dificultad da aprendizaje (DA), utilizando el *Sydney Attribution Scale* para determinar la atribución de los éxitos y los fracasos durante la realización de tareas matemáticas y de lectura a tres posibles causas. Se trata de un diseño con tres factores: un factor intersujetos, *dificultad en el aprendizaje* (con dos niveles: sí; no) y dos factores intrasujetos, *resultados* (con dos niveles: éxito; fracaso) y *causas* (con tres niveles: habilidad; esfuerzo; causas externas). Los resultados

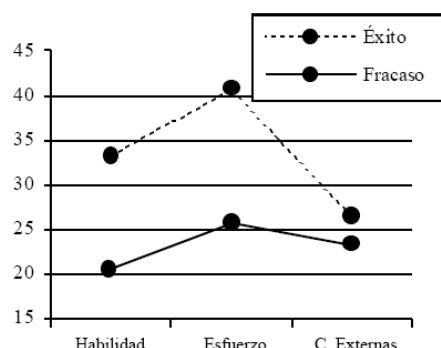
se presentan en las Figuras 1 y 2 del artículo referenciado, que se reproducen en la Figura 4.25. En el apartado resultados (pág. 552) se indica:

“En las Figuras 1 y 2 se muestran gráficamente los patrones atribucionales de cada uno de los dos grupos de estudiantes (DA y NDA, respectivamente). Los análisis de diferencias intragrupo, (prueba «t» para muestras relacionadas) indican que tanto los alumnos con DA como sin DA (NDA) atribuyen sus éxitos significativamente más a su esfuerzo realizado que a su propia habilidad ( $t(1, 222)=15,759$   $p=0,000$  y  $t(1, 210)=13,400$   $p=0,000$ ) ocurriendo lo mismo respecto a la atribución ante resultados negativos ( $t(1, 218)=5,125$   $p=0,000$  y  $t(1, 215)=11,700$   $p=0,000$ , grupos DA y NDA, respectivamente). Por otra parte, mientras que los estudiantes con DA sus éxitos los atribuyen significativamente más a causas externas que a su capacidad,  $t(1, 224)=5,571$   $p=0,000$ , los estudiantes NDA atribuyen los éxitos mucho más a la capacidad que a causas externas  $t(1, 210)=7,472$   $p=0,000$ . En cuanto a los fracasos, mientras que los estudiantes con DA los atribuyen significativamente más a la falta de capacidad que a causas externas  $t(1, 214)=4,117$   $p=0,000$ , los estudiantes NDA los atribuyen más a causas externas que a su falta de capacidad  $t(1, 209)=4,789$   $p=0,000$ .”

**Figura 4.25. Tomada del original (*Psicothema*, 2000)**



**Figura 1.** Atribución causal para resultados Positivos y negativos en estudiantes con DA



**Figura 2.** Atribución causal para resultados Positivos y negativos en estudiantes sin DA

En el párrafo anterior en ningún momento se hace alusión a la posible interacción entre los tres factores estudiados, si bien la comparación de las atribuciones causales entre los grupos DA y NDA en situaciones de éxito y fracaso son el objetivo del estudio. Así en el resumen del artículo (pág. 548) se dice:

“Los resultados del estudio muestran que estudiantes con DA, *en relación a sus iguales sin dificultades*, presentan una imagen más negativa a nivel general...atribuyen sus fracasos más a causas internas que externas y se responsabilizan menos de sus éxitos.” (cursiva mía).

Se afirma que las atribuciones causales (factor 1) que hacen los grupos (factor 2) *no son las mismas frente al éxito que frente al fracaso* (factor 3). Esto equivale a afirmar que existe una interacción triple. Pero tal interacción no se ha analizado formalmente y las comparaciones entre casillas que nos ofrecen los resultados, *efectos simples simples*, no permiten llegar de forma justificada a la conclusión de que tal interacción existe.

**Ejemplo 37)** (*Psicológica*, 2000) Los autores utilizaron el procedimiento de supresión condicionada para estudiar el papel de la asociación contexto-estímulo incondicionado en la inhibición diferencial. Para analizar las pruebas de sumación se utilizó un diseño de cuatro factores: dos factores intersujetos, *grupo [ITI]* (con tres niveles: short-short [S-S]; short-long [S-L]; long-short [L-S]) y *presentación* de estímulos (con dos niveles: serial [S]; random [R]), y dos factores intrasujetos, *sumación* (con dos niveles: L; LT) y *ensayos* (con cuatro niveles: 1º; 2º; 3º; 4º). En el apartado resultados (págs. 224-5) se dice:

“... Interactions occurred for summation  $\times$  presentation  $\times$  trials,  $F(3, 27)=6.87$ , and summation  $\times$  ITI,  $F(2, 29)=3.35$ . *Several simple effects analysis were conducted to explore these interactions.*”

Más adelante en la pág. 225 se dice:

“... A three-way ANOVA for summation in S-S groups was conducted. Summation effects was reliable  $F(1, 9)=8.66$ , as well as summation  $\times$  groups  $\times$  trials interaction,  $F(3, 27)=3.56$ . *A simple effect analysis, keeping presentations and intervals constant,*



*showed effect summation only in S-S(R)  $F(1, 29)=6.99$ . Trial by trial analysis showed summation on 1<sup>st</sup> trial,  $F(1, 14)=16.73$ , 2<sup>nd</sup> trial,  $F(1, 14)=5.97$ , and 3<sup>rd</sup> trial,  $F(1, 14)=4.34$ .*

Los *efectos simples simples*, es decir, la comparación de pares de celdas, en niveles determinados de los otros dos factores, “*keeping presentations and intervals constant*”, es el procedimiento utilizado en ambos análisis para explorar la interacción.

**Ejemplo 38** (*Psicothema*, 2002) Los autores estudiaron la modulación emocional de los reflejos defensivos mediante la visualización de imágenes de diferente contenido afectivo. Para ello estudiaron la respuesta cardiaca de defensa (RCD) durante los 80 segundos posteriores a la provocación de un sobresalto mediante un estímulo auditivo en presencia de diapositivas agradables, desagradables y neutras en 60 sujetos. Cada participante fue sometido a 9 ensayos, tres con cada tipo de diapositiva, que se presentaron en diferente orden balanceado según un procedimiento de cuadrado latino. Los 80 segundos de registro continuo se dividieron en 10 intervalos en cada uno de los cuales se calculó la mediana de la frecuencia cardiaca. Se trata de un diseño con cuatro factores: un factor intersujetos, *orden de presentación* (con tres niveles) y tres factores intrasujetos, *contenido afectivo* (con 3 niveles: agradable, desagradable y neutro), *estímulos* (con 3 niveles: 1<sup>a</sup>, 2<sup>a</sup> ó 3<sup>a</sup> presentación para cada tipo de contenido) y *medianas* (con 10 niveles). Los resultados se presentan en la Figura 1 del artículo, que se reproduce en la Figura 4.26. En el apartado de resultados (pág. 704) se señala:

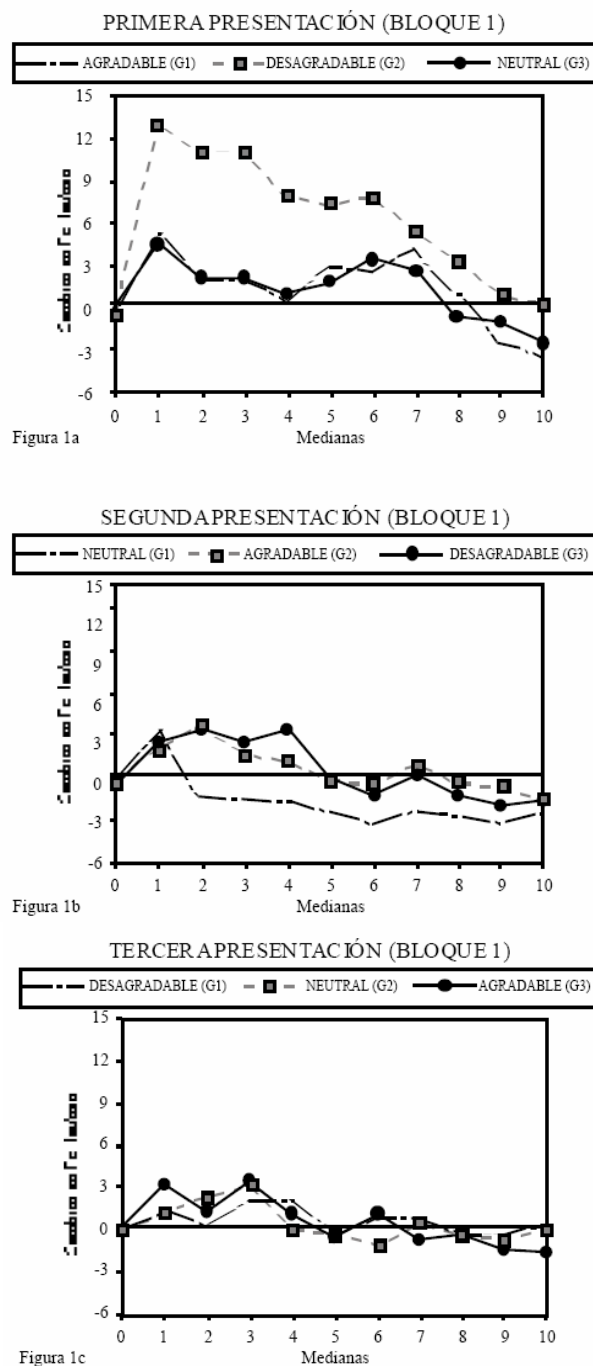
“Los efectos significativos del ANOVA aparecen... [se describen los efectos principales] en las interacciones.... [se describen interacciones de 1<sup>er</sup>, 2<sup>o</sup> y 3<sup>er</sup> orden] Contenido afectivo  $\times$  Medianas  $\times$  Orden ( $F(36, 1026)=1,66$  ( $p=0,0481$ )...

...No obstante, las interacciones significativas encontradas dejan claro que las diferencias en el patrón de la respuesta cardiaca de defensa dependen no sólo del

contenido afectivo de las diapositivas, sino también del orden de presentación y de los estímulos. En la Figura 1 aparecen las representaciones de la RCD para las tres diapositivas (agradable, neutral y desagradable) cuando las tres aparecen en primer lugar.

Los resultados del análisis de cada una de las medianas para las tres presentaciones de las diapositivas del primer bloque muestran diferencias significativas en las medianas 1, 2, 3 y 4 de la primera presentación (ensayo 1). Como puede verse en la Figura 1a, el grupo que ve la diapositiva desagradable en primer lugar tiene mayor frecuencia cardíaca en las medianas 1, 2, 3 y 4 que los otros dos grupos, que ven la diapositiva las diapositivas neutral y agradable también en primer lugar. En la segunda presentación continúa teniendo mayor frecuencia cardíaca el grupo que ve la diapositiva desagradable, aunque las diferencias significativas se reducen a la mediana 4 (ver Figura 1b). En la tercera presentación, las diferencias significativas desaparecen (ver Figura 1c)”

Se está recurriendo a la comparación directa entre pares de celdas, *efectos simples simples*, para explicar una interacción triple. A falta de los análisis complementarios que apoyasen la siguiente explicación, una interpretación razonable de la interacción de segundo orden basada en la observación de la Figura 4.26, podría ser como sigue: “en la primera presentación del bloque 1 (Figura 1a) la diferencia de efecto entre las diapositivas desagradables y las diapositivas agradables (o neutras) sobre la frecuencia cardíaca (variable dependiente) es mayor en las medianas 1 a 4 que en el resto de las medianas, lo cual difiere de lo encontrado en la tercera presentación del bloque 1 (Figura 1c) donde la diferencia de efecto entre las diapositivas con diferente valencia sobre la frecuencia cardíaca no sufre modificaciones en función de las medianas. En cuanto a la segunda presentación del bloque 1 (Figura 1b) la situación es parecida a la que se observa en la tercera presentación, por lo que posiblemente la interacción *valencia de las diapositivas*  $\times$  *mediana* difiera significativamente de la encontrada en la primera presentación.” Esta interpretación se basa exclusivamente en una inspección de los gráficos y no en un análisis formal.

**Figura 4.26.** Tomada del original (*Psicothema*, 2002)**Figura 1.** Respuesta cardiaca de defensa a las tres primeras presentaciones del estímulo auditivo (bloque 1) en función del contenido afectivo de la diapositiva visualizada.

## B.2) Interpretación a partir de los efectos simples

**Ejemplo 39) (Psicothema, 2002)** Los autores estudiaron la modulación emocional de la respuesta cardiaca de defensa (RCD) en situaciones de frustración y/o hostigamiento. Para ello, estudiaron la modificación de la tasa cardiaca y la presión sanguínea en tres ensayos en los que la RCD era desencadenada por un estímulo auditivo intenso. El test de RCD se realizaba después de dos tareas en las que se inducía un estado emocional negativo mediante frustración y/u hostigamiento. La tasa cardiaca y presión sanguínea se monitorizaban continuamente registrando la mediana de los registros obtenidos en 10 periodos de duración creciente durante los 80 segundos siguientes al estímulo acústico, calculando el cambio con referencia a la mediana obtenida en los 15 segundos previos al estímulo acústico. Se trata de un diseño con cuatro factores, con dos factores intersujeto, *grupo* (con 4 niveles: control, hostigamiento, frustración y hostigamiento + frustración) y *sexo* (con 2 niveles: hombre y mujer), y dos factores intrasujeto, *medianas* (con 10 niveles: 2, 5, 9, 14, 20, 27, 34, 44, 57 y 70 s) y ensayos (con 3 niveles: 1<sup>er</sup>, 2<sup>o</sup> y 3<sup>er</sup> ensayo). En el apartado resultados (pág. 452) se dice:

“... mostró efectos significativos en el factor Medianas ( $F(9, 855)=22,18, p<0,0001$ ) y en las interacciones Ensayos  $\times$  Medianas ( $F(18, 1710)= 8,09, p<0,0001$ ) y Ensayos  $\times$  Medianas  $\times$  Sexo ( $F(18, 1710)= 2,23, p<0,01$ ).

... *El análisis de la interacción Ensayos  $\times$  Medianas  $\times$  Sexo mostró diferencias entre hombres y mujeres en las medianas 6<sup>a</sup> ( $F(2, 190)= 3,13, p<0,05$ ) y 10<sup>a</sup> ( $F(2, 190)= 4,51, p<0,01$ ) (cursiva mía)*

Se explica la interacción entre tres factores por medio del análisis de los efectos de un factor (sexo) en el segundo factor (medianas) sin tener en cuenta el tercer factor (ensayos).

**Ejemplo 40)** (*Psicothema*, 2004) Los autores estudiaron la relación entre el efecto de compatibilidad de los distractores y el efecto Simon sobre el tiempo de respuesta (TR) en el procesamiento de información. Además se utilizaron diferentes distancias objeto-distractor (O-E). Todos los sujetos pasaron por todas las condiciones experimentales. Se trata de un diseño factorial con tres factores intrasujetos: *compatibilidad de los distractores* (con tres niveles: compatible; incompatible; neutro), *efecto Simon* (con tres niveles: mismo lado; lado contrario; ambos lados) y *distancia O-D* (con dos niveles: 0,65°; 1,30°). En el apartado de resultados (pág. 287) se dice:

“En cuanto a las interacciones entre los factores considerados, resultaron estadísticamente significativas las siguientes: estímulo acústico  $\times$  compatibilidad de los distractores ( $F_{(4, 44)} = 5.074$ ,  $p < 0.01$ ,  $\eta^2 = 0.316$ ) y *estímulo acústico  $\times$  compatibilidad  $\times$  distancia O-D* ( $F_{(4, 44)} = 7.177$ ,  $p < 0.01$ ,  $\eta^2 = 0.395$ )” (cursiva mía).

“... Se realizó un segundo análisis de varianza de medidas repetidas, reduciendo a dos el número de factores (estímulo acústico y distancia O-D) y comprobando su efecto sobre la variable dependiente utilizada para medir la repercusión de los distractores incompatibles en el procesamiento del objetivo, es decir, el efecto de compatibilidad de respuesta (ECR). La medida de ECR se obtuvo sustrayendo a la media de los TRs obtenidos con los distractores incompatibles la media de los TRs obtenidos bajo la condición compatible, en cada una de las seis combinaciones de los dos niveles ahora considerados (dos distancias y tres modalidades de estímulo acústico)

El análisis de varianza para la variable ECR reveló un efecto principal significativo sólo del factor estímulo acústico ( $F_{(2, 22)} = 4.480$ ,  $p = 0.023$ ,  $\eta^2 = 0.706$ )...

... En este segundo análisis también resultó estadísticamente significativa la interacción estímulo acústico  $\times$  distancia O-D ( $F_{(2, 22)} = 4.572$ ,  $p = 0.022$ ,  $\eta^2 = 0.715$ ). En la figura 3 se observa que la interferencia del estímulo acústico sobre el ECR sólo se produce en la distancia de 1.30°.”

La figura 3 del artículo, sobre la que se interpreta la interacción entre los factores, se reproduce en la Figura 4.27.

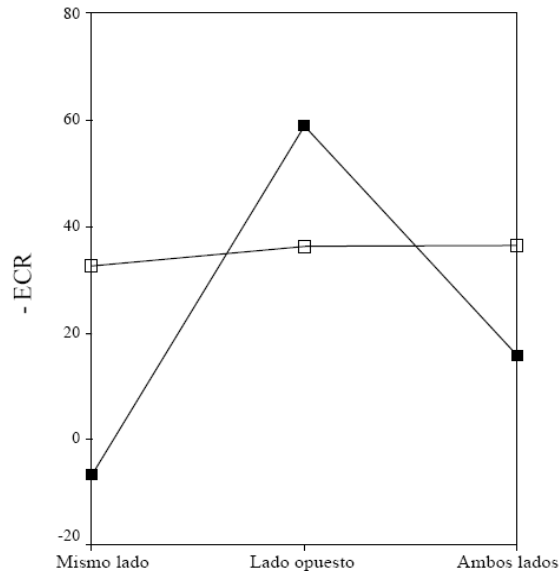
Figura 4.27. Tomada del original (*Psicothema*, 2004)


Figura 3. Efecto de compatibilidad de respuesta (ECR) en función de la procedencia espacial del estímulo acústico (mismo lado, lado opuesto y ambos lados) y de la distancia O-D: (□) 0,65° y (■) 1,30°

El ANOVA de 2 factores, *procedencia espacial del estímulo* (con tres niveles: mismo lado, lado opuesto ambos lados) y *distancia O-D* (con dos niveles: 0,65° y 1,30°), muestra una interacción significativa que no se analiza formalmente sino que se interpreta sobre la figura. En el apartado de discusión (pág. 280) se hace una alusión específica a la interacción entre los tres factores, diciendo:

“La consideración de las modalidades que producen la triple interacción estímulo acústico  $\times$  compatibilidad  $\times$  distancia indica que la interacción entre los dos fenómenos investigados (efecto Simon y ECR) sólo se produjo en la distancia de 1,30°. Por lo tanto, al no darse la interacción entre la distancia menor, podrían plantearse dudas acerca de si efectivamente en todas las condiciones ambos fenómenos afectan o no a una misma etapa de procesamiento.”

En este párrafo la interacción entre tres factores se intenta explicar mediante la interacción encontrada en el ANOVA de dos factores, para lo que se recurre a la interpretación sobre el gráfico de medias.

**Ejemplo 41)** (*Psicothema*, 2005) Los autores estudiaron en 64 voluntarios si los estímulos amenazantes reciben procesamiento automático cuando se presentan enmascarados en un paradigma de *priming*. El *prime* consistió en un rostro amenazante (A) y una cara con expresión neutra como estímulo irrelevante (N). Los mismos estímulos, A y N, se utilizaron como *target*. Los participantes se dividieron en dos grupos en los que la asincronía entre estímulos (SOA) fue 34 ó 51 ms. Finalmente, cada sujeto fue expuesto a 10 ensayos. En cada ensayo, cuando el target aparecía en la pantalla, los sujetos debían pulsar una llave diferente tan rápidamente como fuese posible en función del target mostrado, A o N, siendo la variable dependiente el tiempo de reacción (RT). Se trata de un diseño con 4 factores: 3 factores intrasujetos, *correspondencia* entre prime y target (con dos niveles: congruente; incongruente), *relevancia* (con dos niveles: rostro amenazante [A]; expresión neutra [N]) y *ensayo* (con 10 niveles: 1º a 10º ensayos), y un factor intersujetos, *SOA* (con dos niveles: 34 ms; 51 ms). Parte de los resultados se presentan en la Tabla 1 del artículo que se reproduce en la Tabla 4.14. La interacción entre *SOA*, *correspondencia* y *relevancia* se presenta en el apartado resultados (pág. 334) de la siguiente manera:

“...The interaction between SOA, correspondence of prime/target combinations, and threat-relevance of prime was also significant [ $F(1/62)=6.98$ ,  $p<.05$ ]. The post-hoc test to analyze this interaction revealed that the 34 SOA group exhibited a faster RT in the N/A masking condition than the 51 SOA [ $F(1/62)=7.08$ ,  $p<.01$ ,  $\epsilon=0.62$ ]. The groups did not differ significantly in responding to the other masking condition.”

En el párrafo anterior se comienza hablando de una interacción significativa entre tres factores. Como cada uno de los factores tan sólo tiene dos niveles, se dispone de un único grado de libertad y, por lo tanto, para explicar la interacción basta con facilitar

una interpretación basada en una tabla de medias o gráfico de líneas. Sin embargo, lo que se ha hecho es mezclar dos de las variables implicadas, *congruencia* y *relevancia*, dando lugar a una metavariante, *masking condition*, con 4 niveles resultado de mezclar los niveles de las variables originales para, a continuación, analizar la interacción con un ANOVA 2 (SOA: 34 ms; 51ms)  $\times$  4 (Masking: N/N; A/A; N/A; A/N).

**Tabla 4.14. Tomada del original (*Psicothema*, 2005)**

<p><i>Table 1</i> Mean and standard deviation of the RT data for each masking condition in the 34-ms SOA and 51-ms SOA groups</p>								
SOA	MASKING CONDITIONS							
	N/N		A/A		N/A		A/N	
	M	S.D.	M	S.D.	M	S.D.	M	S.D.
34 ms	655	112	612	79	632	115	677	122
51 ms	660	129	634	105	684	134	687	119

Además, dicho análisis se realiza recurriendo a los efectos simples, como se puede ver en la Tabla 2 del artículo, que se reproduce en la Tabla 4.15.

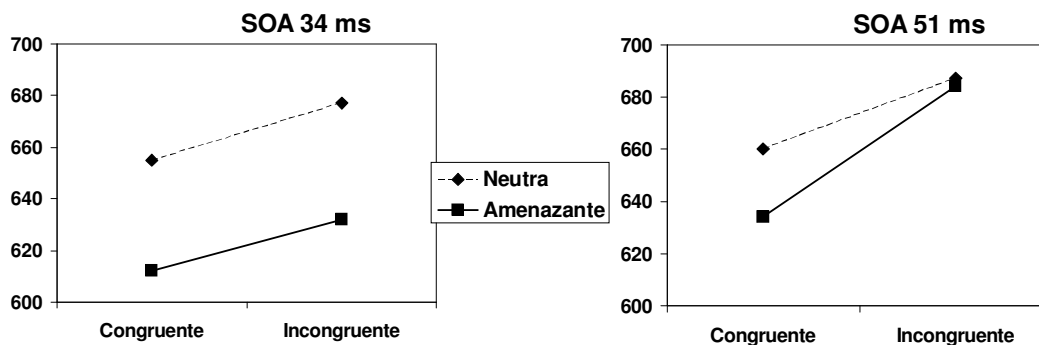
**Tabla 4.15. Tomada del original (*Psicothema*, 2005)**

<p><i>Table 2</i> F-Values obtained in the comparisons of masking conditions in the 34-ms SOA group (in bold) and in the 51-ms SOA group (in italics). *p&lt;.05; ** p&lt;.01</p>				
	N/N	A/A	N/A	A/N
N/N	–	<b>12.35**</b>	<b>3.53</b>	<b>3.25</b>
A/A	<i>6.33*</i>	–	<b>2.67</b>	<b>28.29**</b>
N/A	<i>5.85*</i>	<i>24.36**</i>	–	<b>13.57**</b>
A/N	<i>7.08*</i>	<i>26.82**</i>	<i>0.06</i>	–



La Figura 4.28, preparada a partir de la Tabla 4.14, facilita la interpretación de la interacción de segundo orden. Una interpretación correcta de la interacción significativa entre los tres factores sería: “la interacción *congruencia*  $\times$  *relevancia* se produce de forma diferente en el SOA34 y en el SOA51: mientras en el SOA34 las diferencias en RT entre el *target* neutro y el amenazante son aproximadamente iguales independientemente del nivel de correspondencia, congruente o incongruente, en el SOA51 las diferencias en RT entre el *target* neutro y el amenazante son mayores en la condición congruente que en la condición incongruente.”

Figura 4.28. Preparada a partir de la Tabla 1 del original (*Psicothema*, 2005)



### B.3) Interpretación a partir de los efectos simples de interacción

**Ejemplo 42)** (*Psicothema*, 2001) Los autores exploraron las diferencias entre cambios de tareas regulares y cambios al azar en términos de la magnitud del costo residual asociado al cambio. Se presentaba a los participantes simultáneamente un par de estímulos, un número y una letra, durante 500 ms y se solicitaba que se realizase una de dos tareas posibles: indicar si el número era par o impar, o indicar si la letra era vocal o consonante. El tipo de tarea que había que realizar en cada ensayo era determinado por

la marca de fijación: + para la tarea de números y \* para la tarea de letras, que aparecía 1 s antes (SOA100) o simultáneamente a los estímulos (SOA0). La variable dependiente era el tiempo de reacción (TR). Se trata de un diseño de tres factores intrasujeto: *SOA* (con dos niveles: SOA100; SOA0), *tarea* (con dos niveles: números; letras) y *cambio* (con dos niveles: con cambio de tarea respecto al ensayo anterior [ensayos de cambio]; sin cambio de tarea respecto al ensayo anterior [ensayos de repetición]). Parte de los resultados se muestran en la Figura 1 del artículo reseñado, que se reproduce en la Figura 4.29. En el apartado resultados (pág. 114) se dice:

“...hubo dos interacciones significativas:  $SOA \times Cambio$ ,  $F(1, 17) = 5.415$ ,  $p < 0,033$  y  $SOA \times Tarea \times Cambio$ ,  $F(1, 17) = 6.745$ ,  $p < 0,019$ . *Análisis posteriores indicaron una diferencia clara entre las tareas de números y de letras que daba cuenta de la interacción de segundo orden. En la tarea de letras, el único efecto significativo era el de SOA,  $F(1, 17) = 21.003$ ,  $p < 0,001$ , aunque la variable cambio fue marginalmente significativa,  $F(1, 17) = 3.57$ ,  $p < 0,077$ . La interacción  $SOA \times cambio$  no fue significativa,  $F(1, 17) < 1$ . Para la tarea de números, fueron significativos los efectos principales de SOA,  $F(1, 17) = 21.47$ ,  $p < 0,001$ , y cambio,  $F(1, 17) = 4.887$ ,  $p < 0,042$ , y la interacción  $SOA \times cambio$ ,  $F(1, 17) = 8.11$ ,  $p < 0,012$ ...”*

En el párrafo anterior la interacción entre tres factores se explica con análisis adicionales consistentes en analizar la interacción entre dos factores en cada uno de los niveles del tercero, encontrando que la interacción de dos factores es significativa en uno de ellos y en el otro no. Esto representa una interpretación basada en los efectos simples de interacción que, como ya se ha señalado, no permite dar cuenta del efecto de la interacción. Una interpretación correcta de la interacción triple sería: la interacción *tipo de ensayo*  $\times$  *SOA* en la tarea de letras difiere significativamente de esa misma interacción en la tarea de números; mientras en la tarea de números la diferencia entre el TR medio de los ensayos de cambio y repetición es mayor en el SOA0 que en el

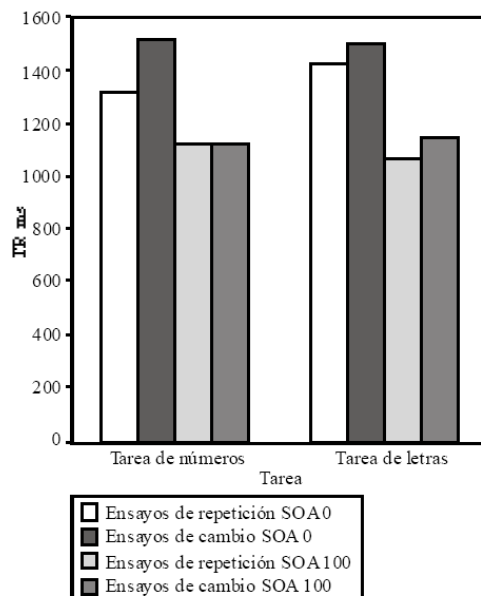
SOA100, en la tarea de letras la diferencia entre el TR medio de los ensayos de cambio y repetición parece ser la misma en el SOA0 que en el SOA100.

El error señalado es el equivalente a explicar la interacción entre dos factores por medio de los efectos simples, siendo al menos uno de ellos significativo y al menos uno de ellos no significativo. Consecuentemente, en un segundo experimento en el que se incluye un nuevo factor, *Predictibilidad*, referido a si los cambios de tarea se producen de forma regular o al azar, los autores hacen la siguiente puntualización (pág. 115):

“...Por el contrario, para los números, la interacción entre el Cambio y la Predictibilidad fue significativo,  $F(1, 38) = 6.632$ ,  $p < 0,015$ . La interacción obedece a que el costo es significativo para el cambio regular,  $F(1, 38) = 9.33$ ,  $p < 0,005$ , pero no lo es para el cambio al azar,  $F(1, 38) < 1$ .

Se está cometiendo el mismo error que antes, al interpretar que una interacción entre dos factores es significativa porque uno de los efectos simples es significativo y el otro no.

**Figura 4.29.** Tomada del original (*Psicothema*, 2001)



*Figura 1. Tiempo de reacción medio para el experimento 1*

***Ejemplo 43*** (IJCHP, 2006) El autor presentó a los sujetos experimentales información sobre el peligroso virus *letrovirus* e información sobre la conducta de detección o prevención de la infección. Cada una de las informaciones podía ser presentada en dos versiones, enfatizando bien la ganancia de seguir la conducta recomendada o bien la pérdida en caso de no seguirla. Después de la lectura, se pidió a los sujetos que contestasen una encuesta sobre su intención conductual y percepción de riesgo, y además se añadió una serie de preguntas que permitía clasificarlos como sujetos con alta o baja necesidad de cognición (NFC) y como sujetos con alta o baja autoestima. Las variables dependientes fueron: *percepción de riesgo* e *intención de realizar la conducta*. Se trata de un modelo con cuatro factores intersujetos: *conducta* presentada (con dos niveles: prevención; detección), *presentación* del mensaje (con dos niveles: ganancia; pérdida), *NFC* (con dos niveles: alta; baja) y *autoestima* (con dos niveles: alta; baja). Respecto a la variable dependiente, *intención de realizar la conducta*, se plantean dos modelos de tres factores. En el apartado de resultados (pág. 621) se indica:

“... En un ANCOVA de  $2(NFC) \times 2(\text{presentación del mensaje}) \times 2(\text{conducta})$ , con la percepción del riesgo para contraer la enfermedad como covariante. Consistente con las predicciones, la interacción  $NFC \times \text{mensaje} \times \text{conducta}$  fue significativa  $F_{(1, 545)} = 6,47$ ;  $p < 0,001$ . Las comparaciones múltiples a posteriori se realizaron por separado para los sujetos altos y bajos en NFC. En los participantes que puntuaron bajo en NFC, la interacción entre presentación del mensaje y conducta resultó significativa,  $F_{(1, 128)} = 9,53$ ;  $p = 0,002$ .... Las intenciones conductuales aportadas por los sujetos que obtuvieron puntuaciones altas NFC no se vieron afectadas por la manipulación del mensaje y de la conducta”.

Es decir, la interacción de segundo orden se interpreta a partir de la interacción entre dos factores en cada nivel del tercero, pero no analizando la diferencia entre ellos, sino indicando que la interacción fue significativa en uno de los niveles del tercer factor, pero no en el otro. Este análisis es innecesario para interpretar la interacción entre los tres factores, pues al tratarse de un diseño  $2 \times 2 \times 2$ , basta con realizar una interpretación adecuada sobre los gráficos de líneas o la tabla de medias.

**Ejemplo 44)** (*Psicothema*, 2004) Los autores estudiaron el efecto conjunto de la intensidad del estímulo incondicionado (EI), descarga eléctrica, y la pre-exposición a un estímulo condicionado (EC), tono acústico, sobre la inhibición de la conducta en ratas entrenadas para obtener refuerzos positivos (bolas de comida) presionando la palanca de una caja de Skinner. Después del entrenamiento y durante el periodo de pre-exposición, la mitad de las ratas fueron expuestas durante 9 sesiones al EC durante periodos de 90s sin presentar el EI, mientras que la otra mitad recibieron sesiones sin EC ni EI. A continuación durante los 6 ensayos de condicionamiento, todas las ratas fueron expuestas al EC de 90s e inmediatamente al EI con un periodo inter-ensayo de 360s. La mitad de las ratas de cada grupo recibieron EI de 0,3 mA y la otra mitad de 0,7 mA. La variable dependiente fue la razón *nº de presiones durante EC / (nº de presiones durante EC + nº de presiones durante 90 s previos)*. Se trata de un diseño con tres factores: dos factores intersujetos, *pre-exposición al EC* (con dos niveles: sí [PE] y no [NPE]) e *intensidad del EI* (con 2 niveles: 0,3 y 0,7 mA) y un factor intrasujetos, *trial* (con 6 niveles: los 6 ensayos de condicionamiento). En la Figura 1 del artículo, reproducida en la Figura 4.30, se presentan los resultados indicando (pág. 210):

“The three-way pre-exposure condition  $\times$  shock intensity  $\times$  trial interaction was significant,  $F(5, 140) = 2,97$ . Subsequent test then demonstrated that the effect of pre-

exposure condition was significant in trial 2,  $F_s(1, 28) = 15,41$ , but not in trials 1, 3, 4, 5 and 6,  $F_s(1, 28) < 1$ . Furthermore, there was a significant pre-exposure condition  $\times$  shock intensity in trials 3, 4 and 5,  $F_s(1, 28) > 6,19$ , but not in Trials 1, 2, and 6,  $F_s(1, 28) < 1$ . Interaction in those trials was due to an effect of pre-exposure on conditioning with a low-intensity shock,  $F_s(1, 14) > 7,12$ , but not with the high-intensity shock  $F_s < 1,18$ ;  $p > 0,29...$

En el párrafo anterior se comentan tanto los efectos simples de pre-exposición en los diferentes ensayos agrupando los niveles de la variable intensidad del shock “the effect of pre-exposure condition was significant in trial 2,  $F_s(1, 28) = 15,41$ , but not in trials 1, 3, 4, 5 and 6,  $F_s(1, 28) < 1$ ”, como los efectos simples de interacción “there was a significant pre-exposure condition  $\times$  shok intensity in trials 3, 4 and 5,  $F_s(1, 28) > 6,19$ , but not in Trials 1, 2, and 6.  $F_s(1, 28) < 1$ .”

Figura 4.30. Tomada del original (Psicothema, 2004)

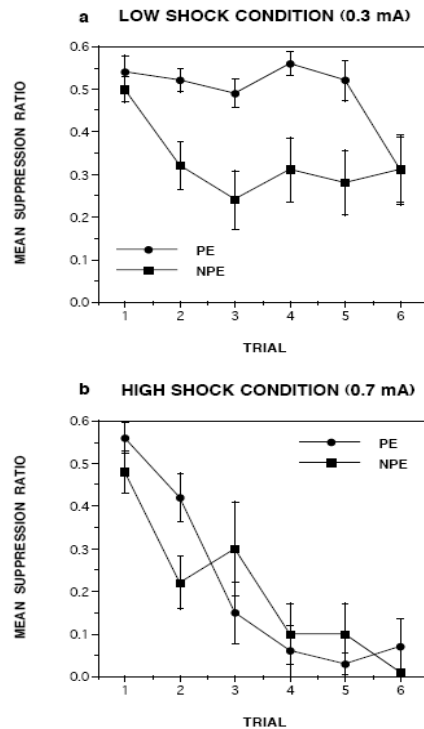


Figure 1. Means and standard errors of suppression ratios during conditioning trials for pre-exposed (PE) and non-pre-exposed (NPE) animals conditioned with low (panel a) or high (panel b) shock intensity.

Un desglose adecuado de la interacción debe incluir los tres factores, indicando por ejemplo que la interacción *pre-exposición*  $\times$  *intensidad del shock* difiere significativamente entre los ensayos 1 y 3, 1 y 4, y 1 y 5 (esta interpretación se basa exclusivamente en la inspección de los gráficos y no en un análisis formal).

### **C) Ejemplos de interpretación correcta de la interacción de segundo orden**

La interacción entre tres factores se interpreta correctamente cuando se basa en la comparación de la interacción entre dos factores entre los niveles del tercero. La interpretación correcta de la interacción de segundo orden tiene diferentes implicaciones en diseños  $2 \times 2 \times 2$  y en diseños en que alguno de los factores tiene más de dos niveles. En el primer caso basta con hacer una interpretación adecuada a partir de la tabla de medias o del gráfico de líneas, ya que la interacción involucra a las ocho celdas disponibles, mientras que en el segundo es necesario realizar análisis adicionales para determinar qué medias están implicadas en la interacción detectada.

En la revisión bibliográfica se han encontrado algunos ejemplos de análisis e interpretación correcta en diseños  $2 \times 2 \times 2$ . Sin embargo, no se ha encontrado ningún artículo en que utilizando diseños de 3 factores con más de dos niveles en alguno de ellos, se hayan realizado análisis adicionales adecuados para determinar qué celdas estaban implicadas en la interacción global encontrada ni, por tanto, se haya realizado una interpretación adecuada.

***Ejemplo 45*** (*Psicothema*, 2003) Los autores estudiaron la atribución diferencial de las dimensiones de moralidad y eficacia al endogrupo y al exogrupo. Para ello, se presentó a los participantes una serie de 48 rasgos relacionados con las dimensiones de moralidad y eficacia, con valencia positiva o negativa (12 en cada combinación dimensión-valencia). Los 24 rasgos con valencia positiva tenían que ser distribuidos para definir el endogrupo y el exogrupo en términos de moralidad y eficacia. La variable dependiente fue el número de rasgos adjudicado a cada categoría definida por un nivel de dimensión, valencia y grupo. Se trata de un diseño con tres factores intersujetos: *valencia del rasgo* (con dos niveles: positivo; negativo), *grupo* (con dos niveles: endogrupo; exogrupo) y *dimensión* (con dos niveles: moralidad; eficacia). Los resultados se presentan en las Figuras 1 y 2 del artículo referenciado, que se reproducen en la Figura 4.31. En el apartado resultados (pág. 410) se dice:

“...La interacción de las tres variables resultó significativa ( $F_{(1, 26)}=56,124$ ;  $p= 0.00$ ). Para aprehender mejor el sentido de este resultado se han dispuesto los datos en dos figuras en función de la variable valencia de rasgo. La Figura 1 representa los resultados obtenidos cuando los rasgos que se tienen que repartir entre el endogrupo y el exogrupo son positivos.

Como se observa en la Figura 1, y tal y como se hipotetizaba, cuando los rasgos son positivos se describe al endogrupo (canarios) más en términos de moralidad ( $M= 7.8$ ,  $D.T.=1.46$ ) que de eficacia ( $M=4.4$ ,  $D.T.=1.73$ ) ( $F_{(1, 26)}=12,10$ ;  $p< 0.01$ ); y al exogrupo (peninsulares) más en términos de eficacia  $M= 7.6$ ,  $D.T.=1.73$ ) que de moralidad ( $M=4.3$ ,  $D.T.=1.48$ ) ( $F_{(1, 26)}=11,36$ ;  $p< 0.01$ ). Por tanto, cuando los rasgos que se tienen que repartir son positivos, se destinan los rasgos de eficacia al exogrupo... y se define el endogrupo en términos de moralidad...

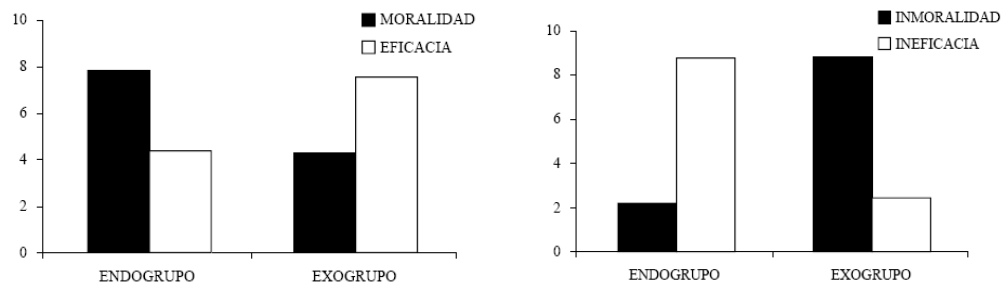
En la Figura 2, que representa los resultados obtenidos cuando los rasgos que tienen que repartirse entre el endogrupo y el exogrupo son negativos, vemos como el patrón de conducta es distinto.



Concretamente, como vemos en la Figura 2, y como se hipotetizaba, cuando los rasgos son negativos, se describe al endogrupo más en términos de eficacia (ineficacia) ( $M=8.8$ ,  $D.T.=2.63$ ) que de moralidad (inmoralidad) ( $M=2.4$ ,  $D.T.=1.67$ ) ( $F_{(1, 26)}=59.44$ ;  $p < 0.001$ ); y al exogrupo más en términos de moralidad (inmoralidad) ( $M=9.0$ ,  $D.T.=1.97$ ) que de eficacia (ineficacia) ( $M=2.8$ ,  $D.T.=1.72$ ) ( $F_{(1, 26)}=52.20$ ;  $p < 0.001$ ). *Esto es, cuando los rasgos que se tienen que repartir entre el endogrupo y el exogrupo son negativos, se invierten los resultados. Ahora se prefiere describir al endogrupo (canarios) en términos de eficacia (ineficacia)... Sin embargo, se define al exogrupo (peninsulares) atendiendo a rasgos de moralidad (inmoralidad)...”(cursiva mía)*

La interpretación de la interacción de segundo orden es correcta, explicando la interacción entre los tres factores mediante el diferente significado de la interacción entre dos factores en cada nivel del tercero. Esta interpretación no requiere de ningún contraste adicional, puesto que se trata de un diseño  $2 \times 2 \times 2$ . No obstante, los investigadores realizan una serie de análisis de los *efectos simples simples*, es decir, contrastes entre dos celdas que utilizan para obtener información adicional.

**Figura 4.31. Tomada del original (Psicothema, 2003)**



**Figura 1.** Media de los rasgos de moralidad y eficacia para el endogrupo (canarios) y el exogrupo (peninsulares) **Figura 2.** Media de los rasgos de inmoralidad e ineficacia para el endogrupo (canario) y el exogrupo (peninsulares)

**Ejemplo 46) (Psicológica, 2006)** Los autores estudiaron la inhibición voluntaria de memorias específicas. Los sujetos se sometieron a un procedimiento en tres fases: aprendizaje de pares asociados, entrenamiento mediante ensayos de “fijar la atención”

(Think) y “no fijar la atención” (No-Think), y tarea de decisión léxica. La variable dependiente era el tiempo de decisión léxica que fue analizado en los ensayos de “fijar la atención” y en los ensayos de “no fijar la atención” por medio de un ANOVA de tres factores intrasujetos: *asincronía* (con dos niveles: SOA100 ms; SOA700 ms), *tipo de ensayo* (con dos niveles: directo [si la pareja *prime-target* había sido aprendida en la fase previa de aprendizaje]; independiente [si dentro de la pareja *prime-target*, el *prime* había sido modificado]) y *cantidad de entrenamiento* (con dos niveles: 0 ensayos; 16 ensayos). Los resultados del experimento 1 se presentan en la Tabla 1 del estudio referenciado, que se reproduce en la Tabla 4.16. En el apartado resultados (pág. 65) se dice:

“Regarding the “Think” data, the analysis of asynchrony (100 vs. 700)  $\times$  type of test (direct vs. independent)  $\times$  training (0 vs. 16) showed significant effects of asynchrony  $F(1, 77) = 11.70$ ,  $Mse=34970.45$ , and type of test,  $F(1, 77) = 14.60$ ,  $Mse=98501.31$ . The interaction of asynchrony by type of test  $F(1, 77) = 18.61$ ,  $Mse=5292.51$ , asynchrony by think condition,  $F(1, 77) = 10.06$ ,  $Mse=4084.60$ , type of test by think condition,  $F(1, 77) = 9.49$ ,  $Mse=2005.86$ , and the higher order interaction of the three variables  $F(1, 77)=11.30$ ,  $Mse=2005.86$ , were also significant. *Looking at the higher order interaction, the analysis indicate that “think” training slower down the decision time to pseudowords at 100 ms at both direct and independent tests. However, at 700 ms, decision time to pseudowords were faster after think training, although the difference disappeared in the independent test...*” (cursiva mía)

En este párrafo se da una interpretación adecuada de la interacción entre los tres factores sin necesidad de recurrir a ningún análisis adicional. Tengase en cuenta que los autores se refieren solo a los ensayos de “fijar la atención” (Think), por lo que se trata de un diseño  $2(asincronía) \times 2(tipo\ de\ ensayo) \times 2(entrenamiento)$  en el que no se requieren contrastes adicionales para explicar la interacción entre tres factores.

**Tabla 4.16.** Tomada del original (*Psicológica*, 2006)

**Table 1.** Mean reaction times, separately by type of analysis (priming, “no think”, and “think”), type of test (direct and Independent) and condition (unrelated vs. related, and amount of training: 0 versus 16) for experiment 1.

SOA	Priming				Think				No Think			
	Direct		Independent		Direct		Independent		Direct		Independent	
	Unrel	Rel	Unrel	Rel	0	16	0	16	0	16	0	16
100	663	641	662	629	695	716	693	711	633	650	632	622
700	627	589	630	584	628	569	662	668	560	592	569	583

## 5. Cómo efectuar contrastes para interpretar la interacción

Para cubrir el segundo objetivo de este trabajo, en los siguientes apartados se explica cómo definir comparaciones lineales entre medias para analizar correctamente una interacción significativa en un diseño factorial con grupos al azar (nos centramos en el caso de 2 y 3 factores) y cómo utilizar el SPSS (programa estadístico más utilizado por los investigadores que publican sus trabajos en las revistas de psicología) para llevar a cabo esas comparaciones.

En los diseños que tienen sólo dos niveles por factor, ya sean diseños  $2 \times 2$  o diseños  $2 \times 2 \times 2$ , cuando se encuentra una interacción significativa no es necesario realizar ninguna otra comparación para interpretarla, pues todas las casillas del diseño están necesariamente involucradas en la interacción. Sin embargo, cuando alguno de los factores tiene más de dos niveles no sucede lo mismo, pues el efecto de la interacción global no indica qué diferencias entre niveles de los factores están involucradas en la interacción (o entre que niveles del tercer factor difieren significativamente las interacciones de primer orden dando lugar a la interacción de segundo orden), lo que debe determinarse por medio de comparaciones adicionales. Esta es una situación que se puede considerar de forma análoga al hallazgo de una  $F$  significativa cuando se están comparando más de dos medias por medio de un ANOVA de un factor.

## 5.1. El caso de dos factores

De la misma manera que es posible definir comparaciones lineales de un grado de libertad para interpretar, descomponiéndolo, un efecto principal (ver, por ejemplo Pardo y San Martín, 1998, págs. 287-292), también es posible definir comparaciones lineales de un grado de libertad para conseguir interpretar una interacción significativa.

Aunque en un diseño factorial el número de comparaciones lineales de un grado de libertad que se pueden plantear es muy elevado (Abelson, 1997, pág. 318), las comparaciones que más ayudan a los investigadores a interpretar una interacción significativa suelen ser aquellas que permiten comparar entre sí los efectos simples.

Cuando un *efecto principal* es declarado significativo se está afirmando que no todas las medias asociadas a ese efecto son iguales. En estos casos, la forma habitual de interpretar lo que está ocurriendo consiste en *comparar por pares* los diferentes niveles del correspondiente factor. Por tanto, para poder interpretar un efecto principal se utilizan comparaciones que involucren dos medias. Y las hipótesis nulas que se plantean son del tipo:

$$H_{0(A)}: \mu_{j+} - \mu_{j'+} = 0 \quad (\text{para todo } j \text{ y } j')$$

Con estas hipótesis se están comparando por pares todas las medias correspondientes a los  $J$  niveles del factor  $A$ . Si el efecto del factor  $A$  es nulo, todas esas diferencias entre medias serán nulas. Si el efecto del factor es significativo, al menos una de esas diferencias será distinta de cero.

Cuando una *interacción de primer orden* es declarada significativa se está afirmando que el efecto del primer factor no es el mismo en todos los niveles del segundo factor. Y, adaptando a esta situación la lógica utilizada con los efectos principales, la

forma de interpretar lo que está ocurriendo consiste en *comparar entre sí los efectos simples* de un factor en los diferentes niveles del otro factor, es decir, en *comparar diferencias*: la diferencia entre dos niveles del primer factor en un nivel del segundo factor con la diferencia entre esos mismos dos niveles del primer factor en un nivel distinto del segundo factor.<sup>9</sup> Por tanto, para interpretar una interacción de primer orden se utilizan comparaciones que involucran cuatro medias. Y las hipótesis nulas que se plantean son del tipo:

$$H_{0(A|B)}: \mu_{jk} - \mu_{j'k} - (\mu_{jk'} - \mu_{j'k'}) = 0 \quad (\text{para todo } j, j', k \text{ y } k')$$

Con estas hipótesis se están comparando entre sí, por pares, los efectos simples del factor A. Si el efecto de la interacción de primer orden es nulo, todas las diferencias entre esos efectos simples serán nulas. Si el efecto de la interacción es significativo, al menos una de esas diferencias entre efectos simples será distinta de cero.

### 5.1.1. Diseños 2×3

En un diseño factorial 2×3 como el propuesto en la Tabla 5.1 es posible interpretar una interacción significativa comparando los efectos simples de A entre los diferentes niveles de B, es decir, comparando entre sí los efectos simples de A (o, si se prefiere, comparando entre sí los efectos simples de B, lo cual es equivalente porque las comparaciones involucradas son las mismas).

---

<sup>9</sup> Cuando el primer factor tiene sólo dos niveles las diferencias entre esos dos niveles del primer factor en cada nivel del segundo factor son los efectos simples propiamente dichos. Cuando el primer factor tiene más de dos niveles estaremos hablando de diferencias simples que no involucran a todos los niveles del primer factor, sino sólo a dos de ellos. Por simplicidad llamaremos a todos ellos efectos simples.

Tabla 5.1. Notación utilizada en un diseño 2×3

	$B_1$	$B_2$	$B_3$
$A_1$	$\mu_{11}$	$\mu_{12}$	$\mu_{13}$
$A_2$	$\mu_{21}$	$\mu_{22}$	$\mu_{23}$

Ahora bien, para comparar entre sí los efectos simples de  $A$  no basta con valorar si un efecto simple es significativo y otro no para, de esta forma, decidir que son distintos. Esto es, básicamente, lo que se hace en los trabajos clasificados en la categoría B de la Tabla 3.1; y ya se ha explicado que esta estrategia no permite aislar el efecto de la interacción porque en los contrastes de los efectos simples se está valorando, además del efecto de la interacción, parte del correspondiente efecto principal (ver ecuación [3.1]). Comparar entre sí los efectos simples de  $A$  requiere:

1. Comparar la diferencia entre  $\mu_{11}$  y  $\mu_{21}$  (o efecto simple de  $A$  en  $B_1$ ) con la diferencia entre  $\mu_{12}$  y  $\mu_{22}$  (o efecto simple de  $A$  en  $B_2$ );
2. Comparar la diferencia entre  $\mu_{11}$  y  $\mu_{21}$  (o efecto simple de  $A$  en  $B_1$ ) con la diferencia entre  $\mu_{13}$  y  $\mu_{23}$  (o efecto simple de  $A$  en  $B_3$ );
3. Comparar la diferencia entre  $\mu_{12}$  y  $\mu_{22}$  (o efecto simple de  $A$  en  $B_2$ ) con la diferencia entre  $\mu_{13}$  y  $\mu_{23}$  (o efecto simple de  $A$  en  $B_3$ ).

Es decir, comparar entre sí los efectos simples de  $A$  requiere efectuar estas tres comparaciones:

$$\psi_1 = A|B_1 - A|B_2 = (\mu_{11} - \mu_{21}) - (\mu_{12} - \mu_{22})$$

$$\psi_2 = A|B_1 - A|B_3 = (\mu_{11} - \mu_{21}) - (\mu_{13} - \mu_{23})$$

$$\psi_3 = A|B_2 - A|B_3 = (\mu_{12} - \mu_{22}) - (\mu_{13} - \mu_{23})$$

Ordenando se obtiene:

$$\psi_1 = \mu_{11} - \mu_{12} - \mu_{21} + \mu_{22}$$

$$\psi_2 = \mu_{11} - \mu_{13} - \mu_{21} + \mu_{23}$$

$$\psi_3 = \mu_{12} - \mu_{13} - \mu_{22} + \mu_{23}$$

Y asignando coeficientes (ver Tabla 5.2):

$$\psi_1 = (1)\mu_{11} + (-1)\mu_{12} + (0)\mu_{13} + (-1)\mu_{21} + (1)\mu_{22} + (0)\mu_{23}$$

$$\psi_2 = (1)\mu_{11} + (0)\mu_{12} + (-1)\mu_{13} + (-1)\mu_{21} + (0)\mu_{22} + (1)\mu_{23}$$

$$\psi_3 = (0)\mu_{11} + (1)\mu_{12} + (-1)\mu_{13} + (0)\mu_{21} + (-1)\mu_{22} + (1)\mu_{23}$$

Por supuesto, una interacción de primer orden puede interpretarse tomando como referencia los efectos simples de cualquiera de los dos factores. De hecho, si en lugar de comparar los efectos simples de  $A$  en cada nivel de  $B$  se comparan los efectos simples de  $B$  en cada nivel de  $A$ , se obtiene idéntico resultado. Ahora bien, mientras que en el caso del factor  $A$  únicamente hay un efecto simple por cada nivel de  $B$  (pues el factor  $A$  sólo tiene dos niveles), en el caso del factor  $B$  hay tres efectos simples por cada nivel de  $A$  (6 efectos simples en total):

$$B_{12}|A_1 = \mu_{11} - \mu_{12}$$

$$B_{13}|A_1 = \mu_{11} - \mu_{13}$$

$$B_{23}|A_1 = \mu_{12} - \mu_{13}$$

$$B_{12}|A_2 = \mu_{21} - \mu_{22}$$

$$B_{13}|A_2 = \mu_{21} - \mu_{23}$$

$$B_{23}|A_2 = \mu_{22} - \mu_{23}$$

Con estos 6 efectos simples es posible hacer 15 comparaciones por pares. Sin embargo, no todas ellas tienen sentido. No olvidemos que el objetivo es poder interpretar la interacción. Y tampoco olvidemos que lo que se está afirmando al decir que una



interacción es significativa es que los efectos simples correspondientes a un factor no son iguales en todos los niveles del otro factor. Por tanto, lo que realmente interesa averiguar es si un efecto simple concreto es o no el mismo en los diferentes niveles del segundo factor. Interesa averiguar, por ejemplo, si el efecto simple asociado a la combinación  $B_{12}$  es o no el mismo en los dos niveles del factor  $A$ ; pero las comparaciones entre los tres efectos simples asociados a  $A_1$  no tienen utilidad (pues no están informando acerca de cómo los efectos simples de  $B$  se comportan en los diferentes niveles de  $A$ ). De acuerdo con este argumento, las comparaciones que interesa efectuar para evaluar el comportamiento de los efectos simples de  $B$  en los diferentes niveles de  $A$  son las siguientes:

$$\psi_1 = B_{12}|A_1 - B_{12}|A_2 = \mu_{11} - \mu_{12} - (\mu_{21} - \mu_{22})$$

$$\psi_2 = B_{13}|A_1 - B_{13}|A_2 = \mu_{11} - \mu_{13} - (\mu_{21} - \mu_{23})$$

$$\psi_3 = B_{23}|A_1 - B_{23}|A_2 = \mu_{12} - \mu_{13} - (\mu_{22} - \mu_{23})$$

Y ocurre que estas tres comparaciones son idénticas a las tres propuestas para comparar entre sí los efectos simples de  $A$  (ver Tabla 5.2).

Por supuesto, estas tres comparaciones son sólo algunas de las que es posible definir para intentar interpretar una interacción significativa, pero tienen más utilidad de la que aparentan. Por ejemplo, no es infrecuente que un investigador esté interesado en analizar una tendencia concreta. Una situación típica de este tipo es aquella en la que la diferencia entre los niveles del factor  $A$  va aumentando (o disminuyendo) conforme progresan los niveles del factor  $B$  (resultado muy típico de los ensayos clínicos en los que se aplica un tratamiento y se hace seguimiento), es decir:

$$\psi_4 = (\mu_{11} - \mu_{21}) < (\mu_{12} - \mu_{22}) < (\mu_{13} - \mu_{23})$$

Esta comparación representa la interacción entre el factor  $A$  y el componente lineal del factor  $B$ , es decir, representa una situación en la que la diferencia entre los niveles del factor  $A$  se va incrementando linealmente a través de los niveles del factor  $B$  (por supuesto, podría trabajarse con un incremento cuadrático utilizando la misma lógica). Los coeficientes que permiten definir esta comparación (ver, por ejemplo, Kirk, 1995, 392-394) se obtienen multiplicando (interacción) los coeficientes que definen la diferencia entre los niveles de  $A$  (coeficientes de  $\psi_A$ ) por los que definen una tendencia lineal en los niveles de  $B$  (coeficientes de  $\psi_{B(\text{lineal})}$ ). La Tabla 5.2 recoge los coeficientes correspondientes a  $\psi_A$  y  $\psi_{B(\text{lineal})}$  (éstos últimos se pueden encontrar en cualquier tabla de polinomios ortogonales), y el producto entre estas dos columnas arroja los coeficientes que recoge la columna  $\psi_{A \times B(\text{lineal})}$ . Ahora bien, estos coeficientes se corresponden con los ya utilizados previamente en la comparación  $\psi_2$  para comparar los efectos simples de  $A$  en  $B_1$  y  $B_3$ . Es decir, la comparación  $\psi_4$  es equivalente a la  $\psi_2$ .

**Tabla 5.2. Coeficientes para comparaciones lineales en un diseño 2x3**

Casillas	$\psi_1$	$\psi_2$	$\psi_3$	$\psi_A$	$\psi_{B(\text{lineal})}$	$\psi_{A \times B(\text{lineal})}$	$\psi_{A B1}$
$A_1 B_1$	1	1	0	1	-1	-1	1
$A_1 B_2$	-1	0	1	1	0	0	0
$A_1 B_3$	0	-1	-1	1	1	1	0
$A_2 B_1$	-1	-1	0	-1	-1	1	-1
$A_2 B_2$	1	0	-1	-1	0	0	0
$A_2 B_3$	0	1	1	-1	1	-1	0

Una característica interesante de las comparaciones  $\psi_1$ ,  $\psi_2$  y  $\psi_3$  es que los coeficientes que las definen permiten comprobar que todas ellas son independientes de los efectos principales (dos comparaciones son independientes si la suma de los productos de sus coeficientes vale cero; ver, por ejemplo, Kirk, 1995, págs. 115-116). Los coeficientes de

la Tabla 5.2 permiten comprobar, por ejemplo, que al sumar los productos de los coeficientes de  $\psi_1$  y  $\psi_A$  se obtiene:

$$(1)(1)+(-1)(1)+(0)(1)+(-1)(-1)+(1)(-1)+(0)(-1) = 0$$

Con los coeficientes de las comparaciones  $\psi_2$  y  $\psi_3$  se obtiene el mismo resultado. Y también se obtiene el mismo resultado si se toma como referente el efecto principal de  $B$ . Sin embargo, esto no ocurre con las comparaciones que definen los efectos simples (ya se ha señalado que un efecto simple incluye tanto parte de la interacción como del efecto principal; ver ecuación [2.5]). Por ejemplo, la columna  $\psi_{A|B_1}$  de la Tabla 5.2 recoge los coeficientes correspondientes a la comparación que define el efecto simple de  $A$  en  $B_1$ . Sumando los productos de estos coeficientes por los que definen el efecto principal de  $A$  se obtiene:

$$(1)(1)+(0)(1)+(0)(1)+(-1)(-1)+(0)(-1)+(0)(-1) = 2$$

Este resultado confirma, por otra vía, la afirmación de la ecuación [3.1] de que los efectos simples no son independientes de los efectos principales.

Según veremos enseguida, el estadístico propuesto más adelante en [5.1] puede utilizarse para evaluar este tipo de comparaciones.

### 5.1.2. Diseños 3×3

Al aumentar los niveles de uno de los dos factores (o de los dos) aumenta rápidamente el número de comparaciones que es necesario diseñar para poder interpretar una interacción significativa. La Tabla 5.3 muestra la notación utilizada en un diseño 3×3.

**Tabla 5.3. Notación utilizada en un diseño 3×3**

	$B_1$	$B_2$	$B_3$
$A_1$	$\mu_{11}$	$\mu_{12}$	$\mu_{13}$
$A_2$	$\mu_{21}$	$\mu_{22}$	$\mu_{23}$
$A_3$	$\mu_{31}$	$\mu_{32}$	$\mu_{33}$

Ahora, cada factor tiene asociados tres efectos simples por cada nivel del otro factor (9 efectos simples en total):

$$A_{12}|B_1 = \mu_{11} - \mu_{21}$$

$$A_{13}|B_1 = \mu_{11} - \mu_{31}$$

$$A_{23}|B_1 = \mu_{21} - \mu_{31}$$

$$A_{12}|B_2 = \mu_{12} - \mu_{22}$$

$$A_{13}|B_2 = \mu_{12} - \mu_{32}$$

$$A_{23}|B_2 = \mu_{22} - \mu_{32}$$

$$A_{12}|B_3 = \mu_{13} - \mu_{23}$$

$$A_{13}|B_3 = \mu_{13} - \mu_{33}$$

$$A_{23}|B_3 = \mu_{23} - \mu_{33}$$

Con estos 9 efectos simples es posible hacer 36 comparaciones por pares. Sin embargo, no todas ellas tienen sentido. El objetivo de estas comparaciones sigue siendo el de poder interpretar la interacción. Y lo que se está afirmando al decir que una interacción es significativa es que los efectos simples correspondientes a un factor no son iguales en todos los niveles del otro factor. Por tanto, lo que realmente interesa averiguar es si un efecto simple concreto es o no el mismo en los diferentes niveles del segundo factor. Interesa averiguar, por ejemplo, si el efecto simple asociado a la combinación  $A_{12}$  es o no el mismo en los niveles del factor  $B$ ; pero las comparaciones entre los tres efectos simples asociados al mismo nivel de  $B$  no tienen utilidad (pues estas comparaciones no informan acerca de cómo los efectos simples de  $A$  se comportan en los diferentes

niveles de  $B$ ). Según esto, las comparaciones que interesa efectuar para evaluar el comportamiento de los efectos simples del factor  $A$  en los diferentes niveles del factor  $B$  son estas 9:

$$\psi_1 = A_{12}|B_1 - A_{12}|B_2 = \mu_{11} - \mu_{21} - (\mu_{12} - \mu_{22})$$

$$\psi_2 = A_{12}|B_1 - A_{12}|B_3 = \mu_{11} - \mu_{21} - (\mu_{13} - \mu_{23})$$

$$\psi_3 = A_{12}|B_2 - A_{12}|B_3 = \mu_{12} - \mu_{22} - (\mu_{13} - \mu_{23})$$

$$\psi_4 = A_{13}|B_1 - A_{13}|B_2 = \mu_{11} - \mu_{31} - (\mu_{12} - \mu_{32})$$

$$\psi_5 = A_{13}|B_1 - A_{13}|B_3 = \mu_{11} - \mu_{31} - (\mu_{13} - \mu_{33})$$

$$\psi_6 = A_{13}|B_2 - A_{13}|B_3 = \mu_{12} - \mu_{32} - (\mu_{13} - \mu_{33})$$

$$\psi_7 = A_{23}|B_1 - A_{23}|B_2 = \mu_{21} - \mu_{31} - (\mu_{22} - \mu_{32})$$

$$\psi_8 = A_{23}|B_1 - A_{23}|B_3 = \mu_{21} - \mu_{31} - (\mu_{23} - \mu_{33})$$

$$\psi_9 = A_{23}|B_2 - A_{23}|B_3 = \mu_{22} - \mu_{32} - (\mu_{23} - \mu_{33})$$

La Tabla 5.4 muestra los coeficientes correspondientes a cada una de estas comparaciones (cuatro medias por comparación). Por supuesto, para comparar los 3 efectos simples de  $B$  en cada nivel de  $A$  es necesario diseñar exactamente las mismas 9 comparaciones.

**Tabla 5.4. Coeficientes para comparar efectos simples en un diseño 3×3**

	$\mu_{11}$	$\mu_{12}$	$\mu_{13}$	$\mu_{21}$	$\mu_{22}$	$\mu_{23}$	$\mu_{31}$	$\mu_{32}$	$\mu_{33}$
$\Psi_1 = A_{12} B_1 - A_{12} B_2$	1	-1	0	-1	1	0	0	0	0
$\Psi_2 = A_{12} B_1 - A_{12} B_3$	1	0	-1	-1	0	1	0	0	0
$\Psi_3 = A_{12} B_2 - A_{12} B_3$	0	1	-1	0	-1	1	0	0	0
$\Psi_4 = A_{13} B_1 - A_{13} B_2$	1	-1	0	0	0	0	-1	1	0
$\Psi_5 = A_{13} B_1 - A_{13} B_3$	1	0	-1	0	0	0	-1	0	1
$\Psi_6 = A_{13} B_2 - A_{13} B_3$	0	1	-1	0	0	0	0	-1	1
$\Psi_7 = A_{23} B_1 - A_{23} B_2$	0	0	0	1	-1	0	-1	1	0
$\Psi_8 = A_{23} B_1 - A_{23} B_3$	0	0	0	1	0	-1	-1	0	1
$\Psi_9 = A_{23} B_2 - A_{23} B_3$	0	0	0	0	1	-1	0	-1	1

### 5.1.3. Diseños 3×4

La Tabla 5.5 muestra la notación utilizada en un diseño 3×4. El factor  $A$  tiene asociados 3 efectos simples por cada nivel del factor  $B$  (12 efectos simples en total). El factor  $B$  tiene asociados 6 efectos simples por cada nivel del factor  $A$  (18 efectos simples en total).

**Tabla 5.5. Notación utilizada en un diseño 3×4**

	$B_1$	$B_2$	$B_3$	$B_4$
$A_1$	$\mu_{11}$	$\mu_{12}$	$\mu_{13}$	$\mu_{14}$
$A_2$	$\mu_{21}$	$\mu_{22}$	$\mu_{23}$	$\mu_{24}$
$A_3$	$\mu_{31}$	$\mu_{32}$	$\mu_{33}$	$\mu_{34}$

Comencemos con los 12 efectos simples asociados al factor  $A$ :

$$A_{12}|B_1 = \mu_{11} - \mu_{21}$$

$$A_{13}|B_1 = \mu_{11} - \mu_{31}$$

$$A_{23}|B_1 = \mu_{21} - \mu_{31}$$

$$A_{12}|B_2 = \mu_{12} - \mu_{22}$$

$$A_{13}|B_2 = \mu_{12} - \mu_{32}$$

$$A_{23}|B_2 = \mu_{22} - \mu_{32}$$

$$A_{12}|B_3 = \mu_{13} - \mu_{23}$$

$$A_{13}|B_3 = \mu_{13} - \mu_{33}$$

$$A_{23}|B_3 = \mu_{23} - \mu_{33}$$

$$A_{12}|B_4 = \mu_{14} - \mu_{24}$$

$$A_{13}|B_4 = \mu_{14} - \mu_{34}$$

$$A_{23}|B_4 = \mu_{24} - \mu_{34}$$

Con estos 12 efectos simples es posible hacer 66 comparaciones por pares. Pero ya hemos argumentado que no todas ellas tienen sentido. El objetivo de estas comparaciones sigue siendo el poder interpretar la interacción. Y lo que se está afirmando al decir que una interacción es significativa es que los efectos simples correspondientes a un factor no son iguales en todos los niveles del otro factor. Por tanto, lo que realmente interesa averiguar es si un efecto simple concreto es o no el mismo en los diferentes niveles del segundo factor. Interesa averiguar, por ejemplo, si el efecto simple asociado a la combinación  $A_{12}$  es o no el mismo en dos niveles distintos del factor  $B$ ; pero las comparaciones entre los tres efectos simples de  $A$  asociados a un mismo nivel de  $B$  no tienen utilidad (pues no están informando acerca de cómo los efectos simples de  $A$  se comportan en los diferentes niveles de  $B$ ). Según esto, las comparaciones que interesa efectuar para evaluar el comportamiento de los efectos simples del factor  $A$  en los diferentes niveles del factor  $B$  son estas 18:

$$\psi_1 = A_{12}|B_1 - A_{12}|B_2 = \mu_{11} - \mu_{21} - (\mu_{12} - \mu_{22})$$

$$\psi_2 = A_{12}|B_1 - A_{12}|B_3 = \mu_{11} - \mu_{21} - (\mu_{13} - \mu_{23})$$

$$\psi_3 = A_{12}|B_1 - A_{12}|B_4 = \mu_{11} - \mu_{21} - (\mu_{14} - \mu_{24})$$

$$\psi_4 = A_{12}|B_2 - A_{12}|B_3 = \mu_{12} - \mu_{22} - (\mu_{13} - \mu_{23})$$

$$\psi_5 = A_{12}|B_2 - A_{12}|B_4 = \mu_{12} - \mu_{22} - (\mu_{14} - \mu_{24})$$

$$\psi_6 = A_{12}|B_3 - A_{12}|B_4 = \mu_{13} - \mu_{23} - (\mu_{14} - \mu_{24})$$

$$\psi_7 = A_{13}|B_1 - A_{13}|B_2 = \mu_{11} - \mu_{31} - (\mu_{12} - \mu_{32})$$

$$\psi_8 = A_{13}|B_1 - A_{13}|B_3 = \mu_{11} - \mu_{31} - (\mu_{13} - \mu_{33})$$

$$\psi_9 = A_{13}|B_1 - A_{13}|B_4 = \mu_{11} - \mu_{31} - (\mu_{14} - \mu_{34})$$

$$\psi_{10} = A_{13}|B_2 - A_{13}|B_3 = \mu_{12} - \mu_{32} - (\mu_{13} - \mu_{33})$$

$$\psi_{11} = A_{13}|B_2 - A_{13}|B_4 = \mu_{12} - \mu_{32} - (\mu_{14} - \mu_{34})$$

$$\psi_{12} = A_{13}|B_3 - A_{13}|B_4 = \mu_{13} - \mu_{33} - (\mu_{14} - \mu_{34})$$

$$\Psi_{13} = A_{23}|B_1 - A_{23}|B_2 = \mu_{21} - \mu_{31} - (\mu_{22} - \mu_{32})$$

$$\Psi_{14} = A_{23}|B_1 - A_{23}|B_3 = \mu_{21} - \mu_{31} - (\mu_{23} - \mu_{33})$$

$$\Psi_{15} = A_{23}|B_1 - A_{23}|B_4 = \mu_{21} - \mu_{31} - (\mu_{24} - \mu_{34})$$

$$\Psi_{16} = A_{23}|B_2 - A_{23}|B_3 = \mu_{22} - \mu_{32} - (\mu_{23} - \mu_{33})$$

$$\Psi_{17} = A_{23}|B_2 - A_{23}|B_4 = \mu_{22} - \mu_{32} - (\mu_{24} - \mu_{34})$$

$$\Psi_{18} = A_{23}|B_3 - A_{23}|B_4 = \mu_{23} - \mu_{33} - (\mu_{24} - \mu_{34})$$

La Tabla 5.6 muestra los coeficientes correspondientes a estas 18 comparaciones (en cada comparación hay involucradas cuatro medias).

**Tabla 5.6. Coeficientes para comparar efectos simples en un diseño 3×4**

	$\mu_{11}$	$\mu_{12}$	$\mu_{13}$	$\mu_{14}$	$\mu_{21}$	$\mu_{22}$	$\mu_{23}$	$\mu_{24}$	$\mu_{31}$	$\mu_{32}$	$\mu_{33}$	$\mu_{34}$
$\Psi_1 = A_{12} B_1 - A_{12} B_2$	1	-1	0	0	-1	1	0	0	0	0	0	0
$\Psi_2 = A_{12} B_1 - A_{12} B_3$	1	0	-1	0	-1	0	1	0	0	0	0	0
$\Psi_3 = A_{12} B_1 - A_{12} B_4$	1	0	0	-1	-1	0	0	1	0	0	0	0
$\Psi_4 = A_{12} B_2 - A_{12} B_3$	0	1	-1	0	0	-1	1	0	0	0	0	0
$\Psi_5 = A_{12} B_2 - A_{12} B_4$	0	1	0	-1	0	-1	0	1	0	0	0	0
$\Psi_6 = A_{12} B_3 - A_{12} B_4$	0	0	1	-1	0	0	-1	1	0	0	0	0
$\Psi_7 = A_{13} B_1 - A_{13} B_2$	1	-1	0	0	0	0	0	0	-1	1	0	0
$\Psi_8 = A_{13} B_1 - A_{13} B_3$	1	0	-1	0	0	0	0	0	-1	0	1	0
$\Psi_9 = A_{13} B_1 - A_{13} B_4$	1	0	0	-1	0	0	0	0	-1	0	0	1
$\Psi_{10} = A_{13} B_2 - A_{13} B_3$	0	1	-1	0	0	0	0	0	0	-1	1	0
$\Psi_{11} = A_{13} B_2 - A_{13} B_4$	0	1	0	-1	0	0	0	0	0	-1	0	1
$\Psi_{12} = A_{13} B_3 - A_{13} B_4$	0	0	1	-1	0	0	0	0	0	0	-1	1
$\Psi_{13} = A_{23} B_1 - A_{23} B_2$	0	0	0	0	1	-1	0	0	-1	1	0	0
$\Psi_{14} = A_{23} B_1 - A_{23} B_3$	0	0	0	0	1	0	-1	0	-1	0	1	0
$\Psi_{15} = A_{23} B_1 - A_{23} B_4$	0	0	0	0	1	0	0	-1	-1	0	0	1
$\Psi_{16} = A_{23} B_2 - A_{23} B_3$	0	0	0	0	0	1	-1	0	0	-1	1	0
$\Psi_{17} = A_{23} B_2 - A_{23} B_4$	0	0	0	0	0	1	0	-1	0	-1	0	1
$\Psi_{18} = A_{23} B_3 - A_{23} B_4$	0	0	0	0	0	0	1	-1	0	0	-1	1

Por supuesto, comparando los efectos simples de  $B$  en cada nivel de  $A$  se llega exactamente a las mismas 18 comparaciones. Ahora bien, mientras que en el caso del factor



A hay 3 efectos simples por cada nivel de  $B$  (es decir, 12 efectos simples en total), en el caso del factor  $B$  hay 6 efectos simples por cada nivel de  $A$  (es decir, 18 efectos simples en total). Estos 18 efectos simples son los siguientes:

$$B_{12}|A_1 = \mu_{11} - \mu_{12}$$

$$B_{13}|A_1 = \mu_{11} - \mu_{13}$$

$$B_{14}|A_1 = \mu_{11} - \mu_{14}$$

$$B_{23}|A_1 = \mu_{12} - \mu_{13}$$

$$B_{24}|A_1 = \mu_{12} - \mu_{14}$$

$$B_{34}|A_1 = \mu_{13} - \mu_{14}$$

$$B_{12}|A_2 = \mu_{21} - \mu_{22}$$

$$B_{13}|A_2 = \mu_{21} - \mu_{23}$$

$$B_{14}|A_2 = \mu_{21} - \mu_{24}$$

$$B_{23}|A_2 = \mu_{22} - \mu_{23}$$

$$B_{24}|A_2 = \mu_{22} - \mu_{24}$$

$$B_{34}|A_2 = \mu_{23} - \mu_{24}$$

$$B_{12}|A_3 = \mu_{31} - \mu_{32}$$

$$B_{13}|A_3 = \mu_{31} - \mu_{33}$$

$$B_{14}|A_3 = \mu_{31} - \mu_{34}$$

$$B_{23}|A_3 = \mu_{32} - \mu_{33}$$

$$B_{24}|A_3 = \mu_{32} - \mu_{34}$$

$$B_{34}|A_3 = \mu_{33} - \mu_{34}$$

Con estos 18 efectos simples es posible hacer 153 comparaciones por pares. Pero no todas ellas tienen sentido. De nuevo lo que interesa averiguar es si un efecto simple concreto es o no el mismo en los diferentes niveles del segundo factor. Interesa averiguar, por ejemplo, si el efecto simple asociado a la combinación  $B_{12}$  es o no el

mismo en niveles distintos del factor  $A$ ; pero las comparaciones entre los seis efectos simples de  $B$  asociados a un mismo nivel de  $A$  no tienen utilidad (pues no están informando acerca de cómo los efectos simples de  $B$  se comportan en los diferentes niveles de  $A$ ). Por tanto, las comparaciones que interesa efectuar para evaluar el comportamiento de los efectos simples del factor  $B$  en los diferentes niveles del factor  $A$  son estas 18:

$$\psi_1 = B_{12}|A_1 - B_{12}|A_2 = \mu_{11} - \mu_{12} - (\mu_{21} - \mu_{22})$$

$$\psi_2 = B_{12}|A_1 - B_{12}|A_3 = \mu_{11} - \mu_{12} - (\mu_{31} - \mu_{32})$$

$$\psi_3 = B_{12}|A_2 - B_{12}|A_3 = \mu_{21} - \mu_{22} - (\mu_{31} - \mu_{32})$$

$$\psi_4 = B_{13}|A_1 - B_{13}|A_2 = \mu_{11} - \mu_{13} - (\mu_{21} - \mu_{23})$$

$$\psi_5 = B_{13}|A_1 - B_{13}|A_3 = \mu_{11} - \mu_{13} - (\mu_{31} - \mu_{33})$$

$$\psi_6 = B_{13}|A_2 - B_{13}|A_3 = \mu_{21} - \mu_{23} - (\mu_{31} - \mu_{33})$$

$$\psi_7 = B_{14}|A_1 - B_{14}|A_2 = \mu_{11} - \mu_{14} - (\mu_{21} - \mu_{24})$$

$$\psi_8 = B_{14}|A_1 - B_{14}|A_3 = \mu_{11} - \mu_{14} - (\mu_{31} - \mu_{34})$$

$$\psi_9 = B_{14}|A_2 - B_{14}|A_3 = \mu_{21} - \mu_{24} - (\mu_{31} - \mu_{34})$$

$$\psi_{10} = B_{23}|A_1 - B_{23}|A_2 = \mu_{12} - \mu_{13} - (\mu_{22} - \mu_{23})$$

$$\psi_{11} = B_{23}|A_1 - B_{23}|A_3 = \mu_{12} - \mu_{13} - (\mu_{32} - \mu_{33})$$

$$\psi_{12} = B_{23}|A_2 - B_{23}|A_3 = \mu_{22} - \mu_{23} - (\mu_{32} - \mu_{33})$$

$$\psi_{13} = B_{24}|A_1 - B_{24}|A_2 = \mu_{12} - \mu_{14} - (\mu_{22} - \mu_{24})$$

$$\psi_{14} = B_{24}|A_1 - B_{24}|A_3 = \mu_{12} - \mu_{14} - (\mu_{32} - \mu_{34})$$

$$\psi_{15} = B_{24}|A_2 - B_{24}|A_3 = \mu_{22} - \mu_{24} - (\mu_{32} - \mu_{34})$$

$$\psi_{16} = B_{34}|A_1 - B_{34}|A_2 = \mu_{13} - \mu_{14} - (\mu_{23} - \mu_{24})$$

$$\psi_{17} = B_{34}|A_1 - B_{34}|A_3 = \mu_{13} - \mu_{14} - (\mu_{33} - \mu_{34})$$

$$\psi_{18} = B_{34}|A_2 - B_{34}|A_3 = \mu_{23} - \mu_{24} - (\mu_{33} - \mu_{34})$$

Y ocurre que estas 18 comparaciones son exactamente las mismas que las 18 propuestas para comparar los efectos simples del factor  $A$  en los diferentes niveles del factor  $B$  (ver Tabla 5.6).

#### 5.1.4. El procedimiento *ANOVA de un factor*

Para contrastar hipótesis del tipo  $\psi_h = 0$  puede utilizarse el estadístico  $t$  convencional (ver Kirk, 1995, págs. 137-143):

$$t = \frac{\hat{\Psi}_h}{\sqrt{MCE \sum_k \frac{c_k^2}{n_k}}} \quad [5.1]$$

aplicando la correspondiente corrección para controlar la tasa de error asociada al número de comparaciones que se lleven a cabo. Este estadístico se distribuye según el modelo de probabilidad  $t$  de *Student* con grados de libertad igual al número de casos menos el número de casillas ( $h$  se refiere a una comparación cualquiera;  $k$  se refiere al número de casillas del diseño;  $MCE$  es la media cuadrática error o error cuadrático medio).

Lo que interesa destacar de esta forma de abordar el estudio de la interacción entre dos factores es que el estadístico propuesto en [5.1] está disponible en el SPSS dentro del procedimiento *ANOVA de un factor*. Ahora bien, para poder utilizar este procedimiento con un diseño de dos factores, es necesario crear una nueva variable cuyos valores sean el resultado de combinar los valores de los factores  $A$  y  $B$ . Por ejemplo, si el factor  $A$  tiene dos niveles (con códigos 1 y 2) y el factor  $B$  tiene 3 (con códigos 1, 2 y 3), la nueva variable deberá tener 6 códigos distintos: 1 para la

combinación  $A_1B_1$ , 2 para  $A_1B_2$ , 3 para  $A_1B_3$ , 4 para  $A_2B_1$ , 5 para  $A_2B_2$  y 6 para  $A_2B_3$ . Una vez creada la nueva variable, basta con introducirla como factor en el mencionado procedimiento y definir en el cuadro de diálogo *Contrastes* las comparaciones que se desea llevar a cabo (ver Pardo y Ruiz, 2005, págs. 354-359, para una explicación de cómo llevar a cabo este tipo de comparaciones).

Por supuesto, esta estrategia es posible utilizarla porque la media cuadrática error (*MCE*) que se obtiene con el modelo factorial original es exactamente la misma que se obtiene cuando se realiza la transformación para aplicar el modelo de un factor (puesto que la *MCE* representa la variabilidad existente dentro de las casillas, su valor no cambia porque se altere el orden de las mismas).

Veamos cómo utilizar el programa estadístico SPSS para efectuar algunas de las comparaciones propuestas en los apartados anteriores (no debemos perder de vista que el objetivo de los ejemplos que se ofrecen a continuación es el de ayudar a interpretar una interacción significativa de primer orden).

Comencemos retomando el *ejemplo 2* ya presentado en la Tabla 3.6. Recordemos que se trataba de un diseño  $2 \times 3$  en el que los efectos principales no resultaron significativos pero sí la interacción (ver Tabla 3.8). El hecho peculiar de este ejemplo es que, a pesar de que la *F* de la interacción es significativa ( $p < 0,05$ ), ninguno de los efectos simples alcanza a serlo ( $p = 0,063$  para los tres efectos simples del factor *A*; ver Tabla 3.9).

Para llevar a cabo las comparaciones que permiten interpretar el efecto de la interacción hay que comenzar creando una nueva variable con códigos distintos para cada combinación *AB*; en concreto, vamos a crear una variable con códigos del 1 al 6:  $1 = A_1B_1$ ,  $2 = A_1B_2$ ,  $3 = A_1B_3$ ,  $4 = A_2B_1$ ,  $5 = A_2B_2$ ,  $6 = A_2B_3$ . La Tabla 5.7 reproduce los datos de la Tabla 3.6 con esta nueva variable (la hemos llamado *C*).

Puesto que nuestro interés radica en descomponer la interacción para poder interpretarla, vamos a realizar las tres comparaciones propuestas para un diseño 2×3 (ver Tabla 5.2):

$$\psi_1 = (1)\mu_{11} + (-1)\mu_{12} + (0)\mu_{13} + (-1)\mu_{21} + (1)\mu_{22} + (0)\mu_{23}$$

$$\psi_2 = (1)\mu_{11} + (0)\mu_{12} + (-1)\mu_{13} + (-1)\mu_{21} + (0)\mu_{22} + (1)\mu_{23}$$

$$\psi_3 = (0)\mu_{11} + (1)\mu_{12} + (-1)\mu_{13} + (0)\mu_{21} + (-1)\mu_{22} + (1)\mu_{23}$$

**Tabla 5.7. Diseño 2×3. Datos del ejemplo 2 con la nueva variable C**

A	B	C	Y	A	B	C	Y	A	B	C	Y
1	1	1	2,0	1	3	3	1,5	2	2	5	2,0
1	1	1	2,4	1	3	3	1,9	2	2	5	2,4
1	1	1	2,6	1	3	3	2,1	2	2	5	2,6
1	1	1	2,8	1	3	3	2,3	2	2	5	2,8
1	1	1	3,0	1	3	3	2,5	2	2	5	3,0
1	1	1	3,0	1	3	3	2,5	2	2	5	3,0
1	1	1	3,2	1	3	3	2,7	2	2	5	3,2
1	1	1	3,4	1	3	3	2,9	2	2	5	3,4
1	1	1	3,6	1	3	3	3,1	2	2	5	3,6
1	1	1	4,0	1	3	3	3,5	2	2	5	4,0
1	2	2	1,5	2	1	4	1,5	2	3	6	2,0
1	2	2	1,9	2	1	4	1,9	2	3	6	2,4
1	2	2	2,1	2	1	4	2,1	2	3	6	2,6
1	2	2	2,3	2	1	4	2,3	2	3	6	2,8
1	2	2	2,5	2	1	4	2,5	2	3	6	3,0
1	2	2	2,5	2	1	4	2,5	2	3	6	3,0
1	2	2	2,7	2	1	4	2,7	2	3	6	3,2
1	2	2	2,9	2	1	4	2,9	2	3	6	3,4
1	2	2	3,1	2	1	4	3,1	2	3	6	3,6
1	2	2	3,5	2	1	4	3,5	2	3	6	4,0

Introduciendo los coeficientes correspondientes a las comparaciones  $\psi_1$ ,  $\psi_2$  y  $\psi_3$  en el subcuadro de diálogo *Contrastes* del procedimiento *ANOVA de un factor* (utilizando la

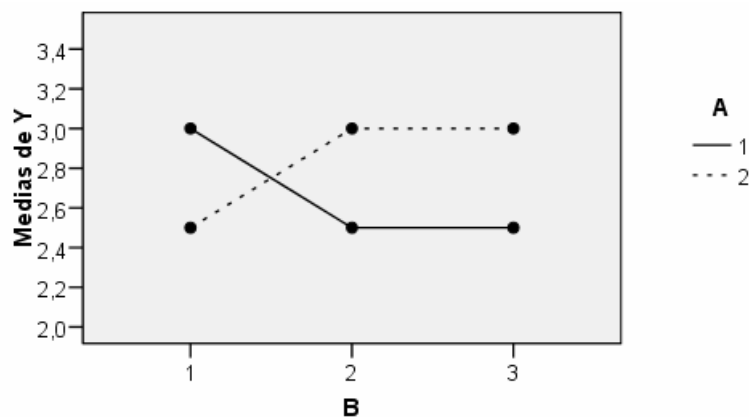
variable  $C$  como variable *factor* y la variable  $Y$  como variable *dependiente*) se obtienen los resultados que muestra la Tabla 5.8:

**Tabla 5. 8. Resultados de las comparaciones  $\psi_1$ ,  $\psi_2$  y  $\psi_3$**

Contraste	Valor del contraste	t	gl	Sig. (bilateral)
$\psi_1 = A B_1 - A B_2$	1,000	2,685	54	<b>0,010</b>
$\psi_2 = A B_1 - A B_3$	1,000	2,685	54	<b>0,010</b>
$\psi_3 = A B_1 - A B_3$	0,000	0,000	54	1,000

Estos resultados tienen un significado claro, perfectamente acorde con la impresión visual que ofrece el gráfico de líneas de la Figura 5.1: el efecto simple de  $A$  en  $B_1$  difiere tanto del efecto simple de  $A$  en  $B_2$  como del efecto simple de  $A$  en  $B_3$ ; y el efecto simple de  $A$  en  $B_2$  no difiere del efecto simple de  $A$  en  $B_3$ . Es decir, lo que ocurre con el efecto de  $A$  en  $B_1$  no es lo mismo que lo que ocurre con el efecto simple de  $A$  en  $B_2$  y  $B_3$ . Y esta afirmación puede hacerse porque se han comparado entre sí los correspondientes efectos simples (cosa que no se hace cuando se valora por separado la significación de cada efecto simple), lo cual confirma que la valoración individual de los efectos simples, sin compararlos entre sí, no es capaz de captar el significado de la interacción.

**Figura 5.1. Diseño 2×3: representación gráfica de las medias de las casillas**



El siguiente ejemplo (que no es más que una ligera modificación del anterior) muestra claramente que mediante el análisis individual de los efectos simples, sin compararlos entre sí, es posible captar erróneamente el significado de la interacción. Se trata de nuevo de un diseño  $2 \times 3$  con dos factores intersujetos  $A$  (con dos niveles:  $A_1, A_2$ ) y  $B$  (con 3 niveles:  $B_1, B_2, B_3$ ). Los datos del ejemplo están recogidos en la Tabla 5.9, la cual incluye, de nuevo, la variable  $C$ .

**Tabla 5.9. Diseño  $2 \times 3$ . Datos del ejemplo 2 modificados**

$A$	$B$	$C$	$X$	$A$	$B$	$C$	$X$	$A$	$B$	$C$	$X$
1	1	1	2,0	1	3	3	1,5	2	2	5	2,0
1	1	1	2,4	1	3	3	1,9	2	2	5	2,4
1	1	1	2,6	1	3	3	2,1	2	2	5	2,6
1	1	1	2,8	1	3	3	2,3	2	2	5	2,8
1	1	1	3,0	1	3	3	2,5	2	2	5	3,0
1	1	1	3,0	1	3	3	2,5	2	2	5	3,0
1	1	1	3,2	1	3	3	2,7	2	2	5	3,2
1	1	1	3,4	1	3	3	2,9	2	2	5	3,4
1	1	1	3,6	1	3	3	3,1	2	2	5	3,6
1	1	1	4,0	1	3	3	3,5	2	2	5	4,0
1	2	2	1,5	2	1	4	1,5	2	3	6	2,2
1	2	2	1,9	2	1	4	1,9	2	3	6	2,6
1	2	2	2,1	2	1	4	2,1	2	3	6	2,8
1	2	2	2,3	2	1	4	2,3	2	3	6	3,0
1	2	2	2,5	2	1	4	2,5	2	3	6	3,2
1	2	2	2,5	2	1	4	2,5	2	3	6	3,2
1	2	2	2,7	2	1	4	2,7	2	3	6	3,4
1	2	2	2,9	2	1	4	2,9	2	3	6	3,6
1	2	2	3,1	2	1	4	3,1	2	3	6	3,8
1	2	2	3,5	2	1	4	3,5	2	3	6	4,2

La Tabla 5.10 muestra las medias correspondientes a cada nivel y a cada combinación de niveles y la Tabla 5.11 ofrece los resultados del ANOVA. Estos resultados indican que los efectos principales no son significativos ( $p = 0,131$  para  $A$ ;  $p = 0,826$  para  $B$ ) pero que sí lo es el efecto de la interacción ( $p = 0,005$ ).

**Tabla 5.10. Medias del diseño 2×3 propuesto en la Tabla 5.9**

	$B_1$	$B_2$	$B_3$	Total
$A_1$	3,00	2,50	2,50	2,67
$A_2$	2,50	3,00	3,20	2,90
Total	2,75	2,75	2,85	2,79

Desv. típ. = 0,59 en todas las casillas

**Tabla 5.11. Tabla resumen del ANOVA**

Fuente	Suma de cuadrados	gl	Media cuadrática	F	Sig.
$A$	0,817	1	0,817	2,356	0,131
$B$	0,133	2	0,067	0,192	0,826
$A*B$	4,133	2	2,067	5,962	<b>0,005</b>
Error	18,720	54	0,347		
Total	23,803	59			

Adoptando la estrategia de interpretar una interacción significativa a partir de la significación individual de los efectos simples se obtienen los resultados que muestra la Tabla 5.12.

**Tabla 5.12. Significación de los efectos simples de  $A$**

	$A_1 - A_2$	$p$
$B_1$	0,500	0,063
$B_2$	-0,500	0,063
$B_3$	-0,700	0,010

Estos resultados indican que el efecto de  $A$  es significativo en el nivel  $B_3$  mientras que no lo es en los niveles  $B_1$  y  $B_2$ . Por tanto, una conclusión basada en estos resultados llevaría a pensar que el efecto del factor  $A$  en el nivel  $B_3$  difiere del efecto del factor  $A$  en los niveles  $B_1$  y  $B_2$ .



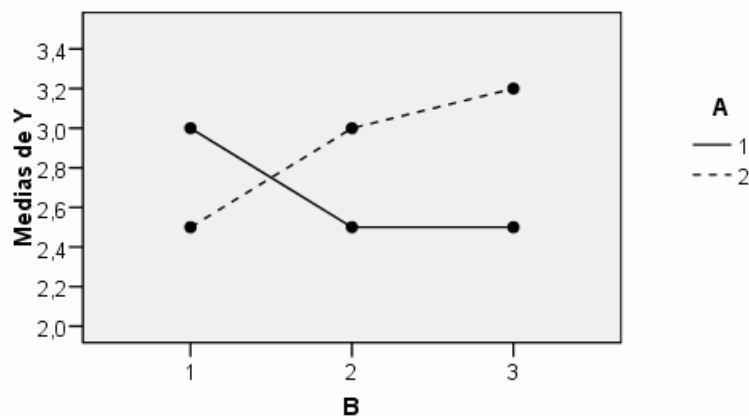
Sin embargo, esto no es lo que realmente está ocurriendo. Utilizando la variable  $C$  en el procedimiento *ANOVA de un factor* del SPSS y asignando coeficientes (los mismos que en el ejemplo anterior) a cada combinación de niveles  $AB$  para definir las comparaciones que permiten interpretar correctamente el efecto de la interacción, se obtienen los resultados que muestra la Tabla 5.13.

**Tabla 5.13. Resultados de las comparaciones  $\psi_1$ ,  $\psi_2$  y  $\psi_3$**

Contraste	Valor del contraste	t	gl	Sig. (bilateral)
$\psi_1 = A B_1 - A B_2$	1,000	2,685	54	<b>0,010</b>
$\psi_2 = A B_1 - A B_3$	1,200	3,223	54	<b>0,002</b>
$\psi_3 = A B_1 - A B_3$	0,200	0,537	54	0,593

El resultado obtenido con esta estrategia tiene un significado bien distinto del obtenido con la valoración individual de la significación de los efectos simples; además, es coherente con la impresión visual que ofrece el gráfico de líneas de la Figura 5.2: el efecto simple de  $A$  en  $B_1$  difiere tanto del efecto simple de  $A$  en  $B_2$  como del efecto simple de  $A$  en  $B_3$ ; y el efecto simple de  $A$  en  $B_2$  no difiere del efecto simple de  $A$  en  $B_3$ . Es decir, lo que ocurre con el efecto de  $A$  en  $B_1$  no es lo mismo que lo que ocurre con el efecto de  $A$  en  $B_2$  y  $B_3$ .

**Figura 5.2. Diseño 2×3. Representación gráfica de las medias**



Es evidente que la conclusión basada en la comparación entre efectos simples difiere de la obtenida previamente al valorar la significación individual de los mismos. Una confusión similar se ha presentado en el *ejemplo 26* (Miñarro y col. 2000), en el que el análisis de los efectos simples lleva a una interpretación diferente de la que se obtendría desglosando el efecto de la interacción mediante comparaciones lineales.

Volvamos a revisar el ejemplo presentado por Howell (2002) con un diseño 2 (*age*: young, old)  $\times$  5 (*condition*: counting, rhyming, adjective, imagery, intention), en el que se obtenían efectos principales significativos para *age* y *condition* y para la interacción *age* $\times$ *condition*. Si, con objeto de interpretar la interacción, se compara el efecto simple de *age* en *rhyming* (una de las tareas de bajo nivel de procesamiento), con los efectos simples de *age* en cada una de las tareas que requieren alto nivel de procesamiento, *adjective*, *imagery* e *intention*, obtenemos los resultados que muestra la Tabla 5.14. Estos resultados indican que el efecto de *age* en *rhyming* no difiere significativamente del efecto de *age* en *adjective* e *imagery*, si bien la diferencia del efecto de *age* en *rhyming* e *intention* si alcanza la significación estadística.

**Tabla 5.14. Algunas comparaciones planeadas para interpretar la interacción del ejemplo presentado por Howell (2002, pág. 425)**

Contraste	Valor del contraste	Sig. (bilateral)
Efecto de <i>age</i> en <i>rhyming</i> vs. efecto de <i>age</i> en <i>adjective</i>	3,1	0,087
Efecto de <i>age</i> en <i>rhyming</i> vs. efecto de <i>age</i> en <i>imagery</i>	3,5	0,054
Efecto de <i>age</i> en <i>rhyming</i> vs. efecto de <i>age</i> en <i>intention</i>	6,6	0,000

De nuevo la interpretación de una interacción significativa a partir de la estrategia basada en la comparación de los efectos simples y a partir de la basada en la significación individual de cada uno de ellos lleva a conclusiones diferentes. ¿Quiere

esto decir que los resultados presentados por Howell no apoyan la teoría de Eysenk? No. Veamos por qué.

Aunque habitualmente las comparaciones que interesa hacer para interpretar una interacción significativa son aquellas que comparan los efectos simples de un factor entre pares de niveles del segundo factor, no siempre éstas son las comparaciones de interés. ¿Cuál o cuáles son las comparaciones que interesa hacer en el experimento de Eysenk? La hipótesis de trabajo de Eysenk es que las diferencias entre grupos de edad sólo se producen en las tareas que requieren un alto nivel de procesamiento. De las cinco tareas realizadas, dos son de bajo nivel de procesamiento (*low level*), mientras que las otras tres son de alto nivel de procesamiento (*high level*). Por tanto, la hipótesis nula que interesa contrastar es que el efecto de la edad en las dos tareas de bajo nivel de procesamiento es igual al efecto de la edad en las tres tareas de alto nivel de procesamiento. Veamos cómo se puede llevarse a cabo este contraste asignando coeficientes a las medias de las casillas. La Tabla 5.15 muestra la notación utilizada en un diseño  $2 \times 5$ .

**Tabla 5.15. Notación utilizada en un diseño  $2 \times 5$**

	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$	
$A_1$	$\mu_{11}$	$\mu_{12}$	$\mu_{13}$	$\mu_{14}$	$\mu_{15}$	$\mu_{1+}$
$A_2$	$\mu_{21}$	$\mu_{22}$	$\mu_{23}$	$\mu_{24}$	$\mu_{25}$	$\mu_{2+}$
	$\mu_{+1}$	$\mu_{+2}$	$\mu_{+3}$	$\mu_{+4}$	$\mu_{+5}$	$\mu$

La hipótesis nula que permite comparar las dos tareas de bajo nivel ( $B_1, B_2$ ) con las tres de alto nivel ( $B_3, B_4, B_5$ ) adopta la forma:

$$H_0 : \frac{(\mu_{11} - \mu_{21}) + (\mu_{12} - \mu_{22})}{2} = \frac{(\mu_{13} - \mu_{23}) + (\mu_{14} - \mu_{24}) + (\mu_{15} - \mu_{25})}{3}$$

Es decir:

$$H_0: 3[(\mu_{11} - \mu_{21}) + (\mu_{12} - \mu_{22})] - 2[(\mu_{13} - \mu_{23}) + (\mu_{14} - \mu_{24}) + (\mu_{15} - \mu_{25})] = 0$$

Asignando coeficientes y ordenando:

$$H_0: (3)\mu_{11} + (3)\mu_{12} + (-2)\mu_{13} + (-2)\mu_{14} + (-2)\mu_{15} \\ + (-3)\mu_{21} + (-3)\mu_{22} + (2)\mu_{23} + (2)\mu_{24} + (2)\mu_{25} = 0$$

La Tabla 5.16 recoge el resultado obtenido al contrastar esta hipótesis. Este resultado es perfectamente acorde con el gráfico de líneas presentado por Howell (ver Figura 3.2) y se debe interpretar de la siguiente manera: el efecto de la edad en el conjunto de las tareas de bajo nivel de procesamiento difiere del efecto de la edad en el conjunto de las tareas de alto nivel de procesamiento.

**Tabla 5.16. Efecto de la edad: diferencia entre las dos tareas de bajo procesamiento y las tres de alto procesamiento (resultados obtenidos a partir de los datos presentados por Howell, 2002, pág. 425)**

	Valor del contraste	Sig. (bilateral)
Efecto de Age en <i>Low level</i> vs. efecto de Age en <i>High level</i>	30,0	0,000

En nuestra opinión, éste es el resultado que verdaderamente apoya la hipótesis de Eysenk, pues se basa en la comparación formal de los efectos de la edad en uno y otro tipo de tareas, comparación que no se lleva a cabo cuando se interpreta la interacción a partir de la significación individual de los efectos simples.

### 5.1.5. La sentencia *LMATRIX*

La estrategia descrita en el apartado anterior es muy sencilla y, por tanto, puede resultar útil para investigadores no demasiado familiarizados con el análisis estadístico. Pero no es la única estrategia disponible para efectuar comparaciones lineales y, desde luego, no es la más elegante. La sentencia *LMATRIX* (disponible mediante sintaxis dentro del procedimiento *UNIANOVA*) también permite contrastar hipótesis del tipo  $\psi_h = 0$ .

Al margen de las comparaciones que es posible llevar a cabo desde el cuadro de diálogo *MODELO LINEAL GENERAL > UNIVARIANTE* (procedimiento *UNIANOVA*) el SPSS ofrece la posibilidad de llevar a cabo cualquier tipo de comparación mediante sintaxis. La sentencia *EMMEANS*, por ejemplo, permite obtener la significación individual de los efectos simples (ver Pardo y Ruiz, 2005, págs. 374-375). Y la sentencia *LMATRIX* permite, entre otras cosas, efectuar comparaciones entre efectos simples.

La sentencia *LMATRIX* permite llevar a cabo cualquier tipo de comparación asignando valores a los coeficientes de la matriz **L** en la hipótesis general **LB = 0** (**B** es el vector de parámetros). El modelo matemático correspondiente a un diseño de dos factores de efectos fijos, completamente al azar, adopta la forma:

$$\mu_{jk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk}$$

(*j* se refiere a los niveles del primer factor y *k* a los niveles del segundo factor). La parte izquierda de la ecuación recoge los pronósticos del modelo. La parte derecha de la ecuación recoge las dos variables independientes y la interacción entre ambas. En un diseño 2×3 el modelo incluye 12 parámetros: la constante, los dos correspondientes a los niveles del factor *A*, los tres correspondientes a los niveles del factor *B* y los seis

correspondientes a la combinación entre los dos niveles del factor  $A$  y los tres del factor  $B$ . Es decir, el vector  $\mathbf{B}$  incluye los siguientes parámetros:

$$\mathbf{B}' = (\mu, \mu_{1+}, \mu_{2+}, \mu_{+1}, \mu_{+2}, \mu_{+3}, \mu_{11}, \mu_{12}, \mu_{13}, \mu_{21}, \mu_{22}, \mu_{23})$$

Y la matriz de coeficientes  $\mathbf{L}$  incluye el peso o coeficiente asignado a cada parámetro del modelo:

$$\mathbf{L} = (l_1, l_2, l_3, l_4, l_5, l_6, l_7, l_8, l_9, l_{10}, l_{11}, l_{12})$$

Para definir contrastes personalizados basta con especificar los valores que deben tomar los coeficientes de la matriz  $\mathbf{L}$  en la expresión  $\mathbf{LB}$ :

$$\begin{aligned} \mathbf{LB} = & l_1\mu + l_2\mu_{1+} + l_3\mu_{2+} + l_4\mu_{+1} + l_5\mu_{+2} + l_6\mu_{+3} + \\ & l_7\mu_{11} + l_8\mu_{12} + l_9\mu_{13} + l_{10}\mu_{21} + l_{11}\mu_{22} + l_{12}\mu_{23} \end{aligned}$$

La sentencia *LMATRIX* permite definir comparaciones de un grado de libertad como las propuestas en los apartados anteriores asignando a cada parámetro los coeficientes apropiados. Así, por ejemplo, para comparar los dos niveles del factor  $A$  en el primer nivel del factor  $B$  (es decir, para evaluar el efecto simple de  $A$  en  $B_1$ ) a los coeficientes  $l_2$  y  $l_7$  asociados a los parámetros correspondientes a la primera categoría de  $A$  ( $\mu_{1+}$ ) y a la combinación de la primera categoría de  $A$  con la primera de  $B$  ( $\mu_{11}$ ) se les asigna un código igual a 1; y a los coeficientes  $l_3$  y  $l_{10}$  asociados a los parámetros correspondientes a la segunda categoría de  $A$  ( $\mu_{2+}$ ) y a la combinación de la segunda categoría de  $A$  con la primera de  $B$  ( $\mu_{21}$ ) se les asigna un código igual a  $-1$ . Al resto de coeficientes se les asignan ceros para excluir del contraste los efectos que no intervienen en la comparación. Por tanto, la expresión  $\mathbf{LB}$  correspondiente a la comparación de los dos niveles del factor  $A$  en el primer nivel de  $B$  queda de la siguiente manera:

$$\mathbf{LB}_{(A1-A2|B1)} = (1)\mu_{1+} + (1)\mu_{11} + (-1)\mu_{2+} + (-1)\mu_{21} = (\mu_{1+} - \mu_{2+}) + (\mu_{11} - \mu_{21})$$

La primera parte de la expresión está comparando los dos niveles del factor  $A$ ; la segunda parte indica que esa comparación se refiere únicamente al primer nivel de  $B$ . De modo similar, la expresión  $\mathbf{LB}$  correspondiente a la comparación entre los dos niveles de  $A$  en el segundo y en el tercer nivel de  $B$  adopta la forma:

$$\mathbf{LB}_{(A1-A2|B2)} = (1)\mu_{1+} + (1)\mu_{12} + (-1)\mu_{2+} + (-1)\mu_{22} = (\mu_{1+} - \mu_{2+}) + (\mu_{12} - \mu_{22})$$

$$\mathbf{LB}_{(A1-A2|B3)} = (1)\mu_{1+} + (1)\mu_{13} + (-1)\mu_{2+} + (-1)\mu_{23} = (\mu_{1+} - \mu_{2+}) + (\mu_{13} - \mu_{23})$$

Ahora bien, comparar el efecto simple de  $A$  en  $B_1$  con el efecto simple de  $A$  en  $B_2$  requiere comparar  $\mathbf{LB}_{(A1-A2|B1)}$  y  $\mathbf{LB}_{(A1-A2|B2)}$ :

$$\begin{aligned}\mathbf{LB}_{(A1-A2|B1)} \text{ y } \mathbf{LB}_{(A1-A2|B2)} &= (\mu_{1+} - \mu_{2+}) + (\mu_{11} - \mu_{21}) - [(\mu_{1+} - \mu_{2+}) + (\mu_{12} - \mu_{22})] \\ &= \mu_{11} - \mu_{21} - \mu_{12} + \mu_{22}\end{aligned}$$

que no es otra cosa que la comparación  $\psi_1$  ya definida para comparar el efecto simple de  $A$  en  $B_1$  con el efecto simple de  $A$  en  $B_2$  (ver Tabla 5.2). Siguiendo el argumento, al comparar los efectos simples definidos en  $\mathbf{LB}_{(A1-A2|B1)}$  y  $\mathbf{LB}_{(A1-A2|B3)}$ , y los definidos en  $\mathbf{LB}_{(A1-A2|B2)}$  y  $\mathbf{LB}_{(A1-A2|B3)}$  se obtienen las comparaciones  $\psi_2$  y  $\psi_3$  (ver Tabla 5.2).

Para llevar a cabo estas comparaciones con el SPSS es necesario añadir, a la sintaxis que se obtiene del cuadro de dialogo *MODELO LINEAL GENERAL > UNIVARIANTE*, la siguiente sentencia *LMATRIX*:

```
/LMATRIX =
a*b  1 -1  0 -1  1  0 ;
a*b  1  0 -1 -1  0  1 ;
a*b  0  1 -1  0 -1  1
```

(repárese en el punto y coma que separa cada comparación). La Tabla 5.17 muestra los resultados que se obtienen con esta estrategia. Puede comprobarse que los valores de los contrastes y de los niveles críticos son idénticos a los obtenidos con la estrategia propuesta en el apartado anterior (ver Tabla 5.8)

**Tabla 5.17. Resultados de las comparaciones  $\psi_1$ ,  $\psi_2$  y  $\psi_3$  con LMATRIX**

Contraste	Valor del contraste	Error típico	Sig. (bilateral)
$L_1 = \psi_1$	1,000	0,372	<b>0,010</b>
$L_2 = \psi_2$	1,000	0,372	<b>0,010</b>
$L_3 = \psi_3$	0,000	0,372	1,000

Un aspecto interesante de esta forma de comparar entre sí los efectos simples es que, además de una valoración individual de cada comparación, se obtiene una valoración global de todas las comparaciones tomadas juntas. La Tabla 5.18 muestra esta valoración. Y ocurre que todos los resultados asociados a este efecto global (suma de cuadrados, grados de libertad, ..., nivel crítico) son idénticos a los del efecto global de la interacción (ver Tabla 3.8). Este resultado indica que las comparaciones que estamos proponiendo para estudiar el efecto de la interacción, no sólo se están refiriendo justamente a ese efecto, sino que lo agotan.

**Tabla 5.18. Efecto global de las comparaciones  $\psi_1$ ,  $\psi_2$  y  $\psi_3$**

Fuente	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Contraste	3,333	2	1,667	4,808	0,012
Error	18,720	54	0,347		



## 5.2. El caso de tres factores

Al igual que en los diseños de dos factores, en los de tres factores también es posible definir comparaciones lineales de un grado de libertad que ayuden a interpretar una interacción significativa. Y, del mismo modo que en los diseños de dos factores las comparaciones que ayudan a entender la interacción de primer orden son las que permiten comparar entre sí los efectos simples, en los diseños de tres factores las comparaciones que ayudan a entender la interacción de segundo orden son, generalmente, las que permiten comparar entre sí los *efectos simples de interacción*.

Cuando un *efecto principal* es declarado significativo se está afirmando que no todas las medias asociadas a ese efecto son iguales. En estos casos, la forma habitual de interpretar lo que está ocurriendo (es decir, la forma de encontrar qué medias difieren de qué otras) consiste en *comparar por pares* los diferentes niveles del correspondiente factor. Por tanto, para poder interpretar un efecto principal se utilizan comparaciones que involucran dos medias. Y las hipótesis nulas que se plantean son del tipo:

$$H_{0(A)}: \mu_{j++} - \mu_{j'++} = 0 \quad (\text{para todo } j \text{ y } j')$$

Con estas hipótesis se están comparando por pares todas las medias correspondientes a los  $J$  niveles del factor  $A$ . Si el efecto del factor  $A$  es nulo, todas esas diferencias entre medias serán nulas. Si el efecto del factor es significativo, al menos una de esas diferencias será distinta de cero.

Cuando una *interacción de primer orden* es declarada significativa se está afirmando que el efecto del primer factor no es el mismo en todos los niveles del segundo factor. Y, adaptando a esta situación la lógica utilizada con los efectos principales, la forma de interpretar lo que está ocurriendo (ver apartado anterior) consiste en *comparar*

entre sí los efectos simples de un factor en los diferentes niveles del otro factor, es decir, en *comparar diferencias*: la diferencia entre dos niveles del primer factor en un nivel del segundo factor con la diferencia entre esos mismos dos niveles del primer factor en un nivel distinto del segundo factor. Por tanto, para interpretar una interacción de primer orden se utilizan comparaciones que involucran cuatro medias. Y las hipótesis nulas que se plantean son del tipo:

$$H_{0(A|B)}: \mu_{jk+} - \mu_{j'k+} - (\mu_{jk'+} - \mu_{j'k'+}) = 0 \quad (\text{para todo } j, j', k \text{ y } k')$$

Con estas hipótesis se están comparando entre sí, por pares, los efectos simples del factor A. Si el efecto de la interacción de primer orden es nulo, todas las diferencias entre esos efectos simples serán nulas. Si el efecto de la interacción es significativo, al menos una de esas diferencias entre efectos simples será distinta de cero.

Cuando una *interacción de segundo orden* es declarada significativa se está afirmando que el efecto de las interacciones de primer orden (el de todas ellas) no es constante (es decir, no es el mismo) en todos los niveles del tercer factor. Y, adaptando a esta situación la lógica utilizada con los efectos principales y con la interacción de primer orden, la forma de interpretar lo que está ocurriendo consiste en comparar entre sí los *efectos simples de interacción*, es decir, en *comparar diferencias de diferencias*: la diferencia de diferencias asociada a la interacción de primer orden en un nivel del tercer factor con la diferencia de diferencias asociada a esa misma interacción de primer orden en un nivel distinto del tercer factor. Por tanto, para interpretar una interacción de segundo orden se utilizan comparaciones que involucran ocho medias. Y las hipótesis nulas que se plantean son del tipo:

$$H_{0(AB|C)}: [\mu_{jkl} - \mu_{j'kl} - (\mu_{jk'l} - \mu_{j'k'l})] - [\mu_{jkl'} - \mu_{j'kl'} - (\mu_{jk'l'} - \mu_{j'k'l'})] = 0 \quad [5.2]$$

(para todo  $j, j', k, k', l$  y  $l'$ )

Con estas hipótesis se están comparando entre sí, por pares, los efectos simples de interacción  $AB|C_l$ .<sup>10</sup> Si el efecto de la interacción de segundo orden es nulo, todas las diferencias entre efectos simples de interacción serán nulas. Si el efecto de la interacción de segundo orden es significativo, al menos una diferencia entre efectos simples de interacción será distinta de cero.

Así las cosas, resulta evidente que no es posible interpretar correctamente una interacción de segundo orden con la estrategia basada en la comparación de dos medias, pues la comparación de dos medias en ningún caso puede captar la complejidad implícita en [5.2] (esta estrategia, denominada, análisis de los *efectos simples simples*, ya ha sido descrita como un error y clasificada en la categoría B1 de la Tabla 3.10). Y tampoco parece posible que pueda interpretarse una interacción de segundo orden agrupando los niveles de uno de los factores y analizando la interacción de primer orden entre los otros dos factores; esta estrategia, obviamente, sólo permite dar cuenta de una interacción de primer orden (también esta estrategia, denominada análisis de los *efectos simples*, ha sido descrita como errónea y clasificada en la categoría B2 de la Tabla 3.10).

Interpretar correctamente una interacción de segundo orden exige comparar entre sí los *efectos simples de interacción*. Ahora bien, comparar entre sí los efectos simples de interacción  $AB|C_l$  no consiste en valorar si la interacción  $AB$  es significativa en un nivel de  $C$  y no en otro para, de esta forma, decidir que la interacción  $AB$  no es la misma en los dos niveles comparados. Esto es, básicamente, lo que se hace en los trabajos clasificados en la categoría B3 de la Tabla 3.10 (*efectos simples de interacción*) y se

---

<sup>10</sup> Cuando los dos primeros factores tienen sólo dos niveles, la diferencia de diferencias entre los dos niveles de un factor en los dos niveles del otro para cada nivel del tercer factor son los efectos simples de interacción propiamente dichos. Cuando el primero y/o el segundo factor tiene más de dos niveles, la diferencia de diferencias entre dos niveles de un factor en dos niveles del otro para cada nivel del tercer factor serán las diferencias simples de diferencias simples que no involucran a todos los niveles del primer factor y/o segundo factor, sino sólo a dos niveles de cada uno de ellos. Por simplicidad llamaremos a todos ellos efectos simples de interacción.

trata de una estrategia errónea equivalente a intentar explicar la interacción entre dos factores a partir de la significación de efectos simples que en realidad no se comparan entre sí. Al explicar los tipos de errores que se cometen al analizar o interpretar una interacción de primer orden (ver capítulo 3) hemos podido comprobar que un efecto simple no informa correctamente sobre la interacción de primer orden porque un efecto simple incluye parte del efecto de la interacción y parte del efecto principal (ver ecuación [3.1]). Ésta es la razón por la cual es necesario comparar entre sí los efectos simples: dado que el efecto principal es el mismo en todos los efectos simples asociados a un factor, la comparación entre efectos simples elimina el efecto principal y se centra exclusivamente en el efecto de la interacción. Con los efectos simples de interacción ocurre algo parecido. Un efecto simple de interacción incluye parte del efecto de la interacción de primer orden y parte del efecto de la interacción de segundo orden:

$$\sum_{l=1}^L SC(AB | C_l) = SCAB + SCABC \quad [5.3]$$

( $SC$  se refiere a las sumas de cuadrados). Por tanto, para eliminar el efecto de la interacción de primer orden y, de esta forma, aislar el efecto de la interacción de segundo orden, es necesario comparar entre sí, por pares, los efectos simples de interacción. Es decir, es necesario poner a prueba hipótesis del tipo  $H_{0(ABC)}$  (ver ecuación [5.3]).

En un diseño  $2 \times 2 \times 2$ , la interacción de segundo orden sólo tiene un grado de libertad. Por tanto, según se ha señalado ya, la interpretación puede hacerse a partir de la tabla de medias o del gráfico de líneas. Cuando alguno de los factores tiene más de dos niveles es necesario hacer algo más.

### 5.2.1. Diseños 2×2×3

En un diseño factorial 2×2×3 como el propuesto en la Tabla 5.19, interpretar una interacción significativa de segundo orden requiere comparar entre sí, por pares, los tres efectos simples de interacción  $AB|C_l$  (o, si se prefiere, comparar entre sí, por pares, los efectos simples de interacción  $AC|B_k$ , o los efectos simples de interacción  $BC|A_j$ , todo lo cual, según veremos, es equivalente).

**Tabla 5.19. Notación utilizada en un diseño 2×2×3**

	$B_1$			$B_2$		
	$C_1$	$C_2$	$C_3$	$C_1$	$C_2$	$C_3$
$A_1$	$\mu_{111}$	$\mu_{112}$	$\mu_{113}$	$\mu_{121}$	$\mu_{122}$	$\mu_{123}$
$A_2$	$\mu_{211}$	$\mu_{212}$	$\mu_{213}$	$\mu_{221}$	$\mu_{222}$	$\mu_{223}$

Ahora bien, puesto que un efecto simple de interacción se define como

$$AB|C_l = \mu_{jkl} - \mu_{j'kl} - (\mu_{jkl} - \mu_{j'kl}) \quad [5.4]$$

los efectos simples de interacción  $AB$  en cada nivel de  $C$  (teniendo en cuenta que  $J=2$ ,  $K=2$  y  $L=3$ ) vendrán dados por:

$$\begin{aligned} AB|C_1 &= \mu_{111} - \mu_{211} - (\mu_{121} - \mu_{221}) \\ AB|C_2 &= \mu_{112} - \mu_{212} - (\mu_{122} - \mu_{222}) \\ AB|C_3 &= \mu_{113} - \mu_{213} - (\mu_{123} - \mu_{223}), \end{aligned} \quad [5.5]$$

Por tanto, interpretar la interacción de segundo orden  $ABC$  a partir de la interacción de primer orden  $AB$  requiere plantear el siguiente grupo de comparaciones:

$$\begin{aligned} \psi_1 &= AB|C_1 - AB|C_2 = \mu_{111} - \mu_{211} - (\mu_{121} - \mu_{221}) - [\mu_{112} - \mu_{212} - (\mu_{122} - \mu_{222})] \\ \psi_2 &= AB|C_1 - AB|C_3 = \mu_{111} - \mu_{211} - (\mu_{121} - \mu_{221}) - [\mu_{113} - \mu_{213} - (\mu_{123} - \mu_{223})] \\ \psi_3 &= AB|C_2 - AB|C_3 = \mu_{112} - \mu_{212} - (\mu_{122} - \mu_{222}) - [\mu_{113} - \mu_{213} - (\mu_{123} - \mu_{223})] \end{aligned}$$

Es decir, requiere plantear las tres comparaciones por pares que es posible hacer con tres efectos. Ordenando se obtiene:

$$\begin{aligned}\Psi_1 &= \mu_{111} - \mu_{112} - \mu_{121} + \mu_{122} - \mu_{211} + \mu_{212} + \mu_{221} - \mu_{222} \\ \Psi_2 &= \mu_{111} - \mu_{113} - \mu_{121} + \mu_{123} - \mu_{211} + \mu_{213} + \mu_{221} - \mu_{223} \\ \Psi_3 &= \mu_{112} - \mu_{113} - \mu_{122} + \mu_{123} - \mu_{212} + \mu_{213} + \mu_{222} - \mu_{223}\end{aligned}\quad [5.6]$$

(por supuesto, estas tres comparaciones no son independientes entre sí, pero eso no es problema; tampoco son independientes entre sí las comparaciones por pares que se hacen *a posteriori* para estudiar un efecto principal y, sin embargo, son las comparaciones que habitualmente interesa hacer). La Tabla 5.20 muestra los coeficientes correspondientes a las tres comparaciones propuestas en [5.6].

**Tabla 5.20. Coeficientes para efectuar comparaciones lineales en un diseño 2×2×3**

	$\mu_{111}$	$\mu_{112}$	$\mu_{113}$	$\mu_{121}$	$\mu_{122}$	$\mu_{123}$	$\mu_{211}$	$\mu_{212}$	$\mu_{213}$	$\mu_{221}$	$\mu_{222}$	$\mu_{223}$
$\Psi_1$	1	-1	0	-1	1	0	-1	1	0	1	-1	0
$\Psi_2$	1	0	-1	-1	0	1	-1	0	1	1	0	-1
$\Psi_3$	0	1	-1	0	-1	1	0	-1	1	0	1	-1
$\Psi_4$	1	0	1	1	0	-1	1	0	-1	-1	0	1
$\Psi_{AB C1}$	1	0	0	-1	0	0	-1	0	0	1	0	0
$\Psi_{A C1}$	1	0	0	1	0	0	-1	0	0	-1	0	0
$\Psi_{A B1C1}$	1	0	0	0	0	0	-1	0	0	0	0	0
$\Psi_A$	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1

$$\Psi_1 = AB|C_1 - AB|C_2 = AC_{12}|B_1 - AC_{12}|B_2 = BC_{12}|A_1 - BC_{12}|A_2$$

$$\Psi_2 = AB|C_1 - AB|C_3 = AC_{13}|B_1 - AC_{13}|B_2 = BC_{13}|A_1 - BC_{13}|A_2$$

$$\Psi_3 = AB|C_2 - AB|C_3 = AC_{23}|B_1 - AC_{23}|B_2 = BC_{23}|A_1 - BC_{23}|A_2$$

Es importante señalar que la interacción de segundo orden puede interpretarse tomando como referencia cualquiera de las tres interacciones de primer orden. De hecho, se obtiene idéntico resultado si en lugar de tomar como referencia la interacción *AB*, se toma como referencia la interacción *AC* o la interacción *BC*.

Si se prefiere trabajar, por ejemplo, con la interacción  $AC$ , hay que comenzar definiendo los efectos simples de interacción asociados a ella. Ahora bien, cuando dos factores tienen sólo dos niveles, como en el caso de la interacción  $AB$ , únicamente hay involucradas cuatro medias por cada nivel del tercer factor:  $\mu_{11l}$ ,  $\mu_{12l}$ ,  $\mu_{21l}$  y  $\mu_{22l}$ . En este caso, todos los efectos simples de interacción asociados a la combinación  $AB$  necesariamente tienen que ser definidos a partir de esas cuatro medias. Esto significa que los efectos simples de interacción  $AB$  son únicos en cada nivel de  $C$  (ver ecuación [5.5]).

Pero éste no es el caso cuando un factor tiene dos niveles y el otro factor tiene tres, como ocurre con la interacción  $AC$ . Con este número de niveles, en la interacción  $AC$  hay involucradas seis medias por cada nivel de  $B$ :  $\mu_{1k1}$ ,  $\mu_{1k2}$ ,  $\mu_{1k3}$ ,  $\mu_{2k1}$ ,  $\mu_{2k2}$  y  $\mu_{2k3}$ . Y como un efecto simple de interacción se construye con cuatro medias (ver ecuación [5.4]), la interacción  $AC$  tiene asociados 3 efectos simples de interacción en cada nivel de  $B$  (6 efectos en total, pues el factor  $B$  tiene 2 niveles):

$$\begin{aligned}
 A_{12}C_{12}|B_1 &= \mu_{111} - \mu_{211} - (\mu_{112} - \mu_{212}) \\
 A_{12}C_{13}|B_1 &= \mu_{111} - \mu_{211} - (\mu_{113} - \mu_{213}) \\
 A_{12}C_{23}|B_1 &= \mu_{112} - \mu_{212} - (\mu_{113} - \mu_{213}) \\
 A_{12}C_{12}|B_2 &= \mu_{121} - \mu_{221} - (\mu_{122} - \mu_{222}) \\
 A_{12}C_{13}|B_2 &= \mu_{121} - \mu_{221} - (\mu_{123} - \mu_{223}) \\
 A_{12}C_{23}|B_2 &= \mu_{122} - \mu_{222} - (\mu_{123} - \mu_{223})
 \end{aligned} \tag{5.7}$$

Con estos 6 efectos simples de interacción es posible hacer 15 comparaciones por pares. Sin embargo, no todas ellas tienen sentido. No olvidemos que el objetivo es poder interpretar la interacción de segundo orden. Y tampoco olvidemos que lo que se está afirmando al decir que una interacción de segundo orden es significativa es que los efectos simples de interacción entre cualesquiera dos factores no son iguales en todos los niveles del tercer factor. Por tanto, lo que realmente interesa averiguar es si un efecto simple

de interacción concreto es o no el mismo en los diferentes niveles del tercer factor. Interesa averiguar, por ejemplo, si el efecto simple asociado a la combinación  $A_{12}C_{12}$  es o no el mismo en los dos niveles del factor  $B$ ; pero las comparaciones entre los tres efectos simples asociados a  $B_1$  no tienen utilidad (pues no están informando acerca de cómo los efectos simples de la interacción  $AC$  se comportan en los diferentes niveles de  $B$ ). Según este argumento, las comparaciones que interesa efectuar para evaluar el comportamiento de la interacción  $AC$  en los diferentes niveles del factor  $B$  son las siguientes:

$$A_{12}C_{12}|B_1 - A_{12}C_{12}|B_2 = \mu_{111} - \mu_{211} - (\mu_{112} - \mu_{212}) - [\mu_{121} - \mu_{221} - (\mu_{122} - \mu_{222})]$$

$$A_{12}C_{13}|B_1 - A_{12}C_{13}|B_2 = \mu_{111} - \mu_{211} - (\mu_{113} - \mu_{213}) - [\mu_{121} - \mu_{221} - (\mu_{123} - \mu_{223})]$$

$$A_{12}C_{23}|B_1 - A_{12}C_{23}|B_2 = \mu_{112} - \mu_{212} - (\mu_{113} - \mu_{213}) - [\mu_{122} - \mu_{222} - (\mu_{123} - \mu_{223})]$$

Y ocurre que estas tres comparaciones son idénticas a las propuestas para comparar entre sí, por pares, los efectos simples asociados a la interacción  $AB$  (ver ecuación [5.6] y las tres primeras comparaciones de la Tabla 5.20).

Por último, si se prefiere utilizar la interacción  $BC$  como referente para la interpretación, puede comprobarse que las comparaciones que es necesario diseñar son exactamente las mismas que en el caso de la interacción  $AB$  y en el de la interacción  $AC$ .

Por supuesto, además de las tres comparaciones propuestas entre efectos simples de interacción, también es posible definir otro tipo de comparaciones para intentar interpretar una interacción significativa (por ejemplo, involucrando más de dos niveles de algún factor, analizando una tendencia, etc.). Así, en un diseño  $2 \times 2 \times 3$  como el que venimos utilizando puede interesar averiguar si el efecto de la interacción  $AB$  va aumentando (o disminuyendo) conforme progresan los niveles del factor  $C$ , es decir:

$$\begin{aligned} \Psi_4 = \Psi_{A \times B \times C(\text{lineal})} &= [\mu_{111} - \mu_{211} - (\mu_{121} - \mu_{221})] < [\mu_{112} - \mu_{212} - (\mu_{122} - \mu_{222})] \\ &< [\mu_{113} - \mu_{213} - (\mu_{123} - \mu_{223})] \end{aligned}$$



Esta comparación representa una situación en la que el efecto de la interacción  $AB$  se va incrementando linealmente a través de los niveles del factor  $C$  (por supuesto, podría estudiarse un incremento cuadrático utilizando la misma lógica). La Tabla 5.20 ofrece los coeficientes correspondientes a esta comparación.

La Tabla 5.20 también ofrece los coeficientes correspondientes a otras comparaciones mencionadas en este apartado, en concreto: el efecto simple de interacción de  $AB$  en  $C_1$  ( $\psi_{AB|C_1}$ ), el efecto simple de  $A$  en  $C_1$  ( $\psi_{A|C_1}$ ), el efecto simple simple de  $A$  en  $B_1C_1$  ( $\psi_{A|B_1C_1}$ ) y el efecto principal de  $A$  ( $\psi_A$ ).

Al igual que sucede con las comparaciones diseñadas para analizar una interacción de primer orden (ver apartado anterior), las comparaciones  $\psi_1$ ,  $\psi_2$  y  $\psi_3$  (todas las cuales incluyen algún componente de la interacción  $AB$ ) son independientes de los efectos principales de  $A$  y de  $B$ , cosa que no ocurre, por ejemplo, ni con los efectos simples ni con los efectos simples simples. Así, si se multiplican los coeficientes de la comparación  $\psi_A$  (efecto principal de  $A$ ; por supuesto, podrían utilizarse, con el mismo resultado, los coeficientes correspondientes al efecto principal de  $B$ ) por los de la comparación  $\psi_1$ , la suma de los productos vale cero. Y lo mismo ocurre si se multiplican los coeficientes de la comparación  $\psi_A$  por los de las comparaciones  $\psi_2$  o  $\psi_3$ . Sin embargo, tanto si los coeficientes de la comparación  $\psi_A$  se multiplican por los de la comparación  $\psi_{A|C_1}$  como si se multiplican por los de la comparación  $\psi_{A|B_1C_1}$ , las sumas de los productos resultantes toman valores distintos de cero (4 en el primer caso y 2 en el segundo).

### 5.2.2. Diseños 2×3×3

En diseños de mayores dimensiones aumenta rápidamente el número de comparaciones que es posible diseñar y, con ello, el de las que es necesario llevar a cabo para entender

el significado de una interacción de segundo orden. La Tabla 5.21 muestra las medias de las casillas en un diseño  $2 \times 3 \times 3$ .

**Tabla 5.21. Notación utilizada en un diseño  $2 \times 3 \times 3$**

	$B_1$			$B_2$			$B_3$		
	$C_1$	$C_2$	$C_3$	$C_1$	$C_2$	$C_3$	$C_1$	$C_2$	$C_3$
$A_1$	$\mu_{111}$	$\mu_{112}$	$\mu_{113}$	$\mu_{121}$	$\mu_{122}$	$\mu_{123}$	$\mu_{131}$	$\mu_{132}$	$\mu_{133}$
$A_2$	$\mu_{211}$	$\mu_{212}$	$\mu_{213}$	$\mu_{221}$	$\mu_{222}$	$\mu_{223}$	$\mu_{231}$	$\mu_{232}$	$\mu_{233}$

En este nuevo escenario, los efectos simples de la interacción  $AB$  siguen siendo tres (uno por cada nivel del factor  $C$ ), pero la interacción  $AB$  ya no es única en cada nivel del factor  $C$ . Ya hemos discutido esta cuestión: si los factores  $A$  y  $B$  tienen sólo dos niveles ( $J = 2$ ;  $K = 2$ ), en la interacción  $AB$  únicamente hay involucradas cuatro medias por cada nivel de  $C$  ( $\mu_{11l}$ ,  $\mu_{12l}$ ,  $\mu_{21l}$  y  $\mu_{22l}$ ) y todos los efectos simples de interacción  $AB$  tienen que ser definidos, necesariamente, a partir de esas cuatro medias (ver ecuación [5.4]). Esto significa que los efectos simples de interacción  $AB$  son únicos en cada nivel de  $C$ :

$$\begin{aligned}
 AB|C_1 &= \mu_{111} - \mu_{211} - (\mu_{121} - \mu_{221}) \\
 AB|C_2 &= \mu_{112} - \mu_{212} - (\mu_{122} - \mu_{222}) \\
 AB|C_3 &= \mu_{113} - \mu_{213} - (\mu_{123} - \mu_{223})
 \end{aligned}
 \tag{5.8}$$

Pero éste no es el caso cuando el factor  $A$  tiene dos niveles y el factor  $B$  tiene tres ( $J = 2$ ;  $K = 3$ ), como en el diseño que nos ocupa. Con este número de niveles, en la interacción  $AB$  hay involucradas seis medias por cada nivel del factor  $C$  ( $\mu_{11l}$ ,  $\mu_{12l}$ ,  $\mu_{13l}$ ,  $\mu_{21l}$ ,  $\mu_{22l}$  y  $\mu_{23l}$ ). Y como un efecto simple de interacción se construye con cuatro medias (ver ecuación [5.4]), la interacción  $AB$  tiene asociados 3 efectos simples de interacción en cada nivel del factor  $C$  (es decir, 9 efectos en total):

$$\begin{aligned}
 A_{12}B_{12}|C_1 &= \mu_{111} - \mu_{211} - (\mu_{121} - \mu_{221}) \\
 A_{12}B_{13}|C_1 &= \mu_{111} - \mu_{211} - (\mu_{131} - \mu_{231})
 \end{aligned}$$

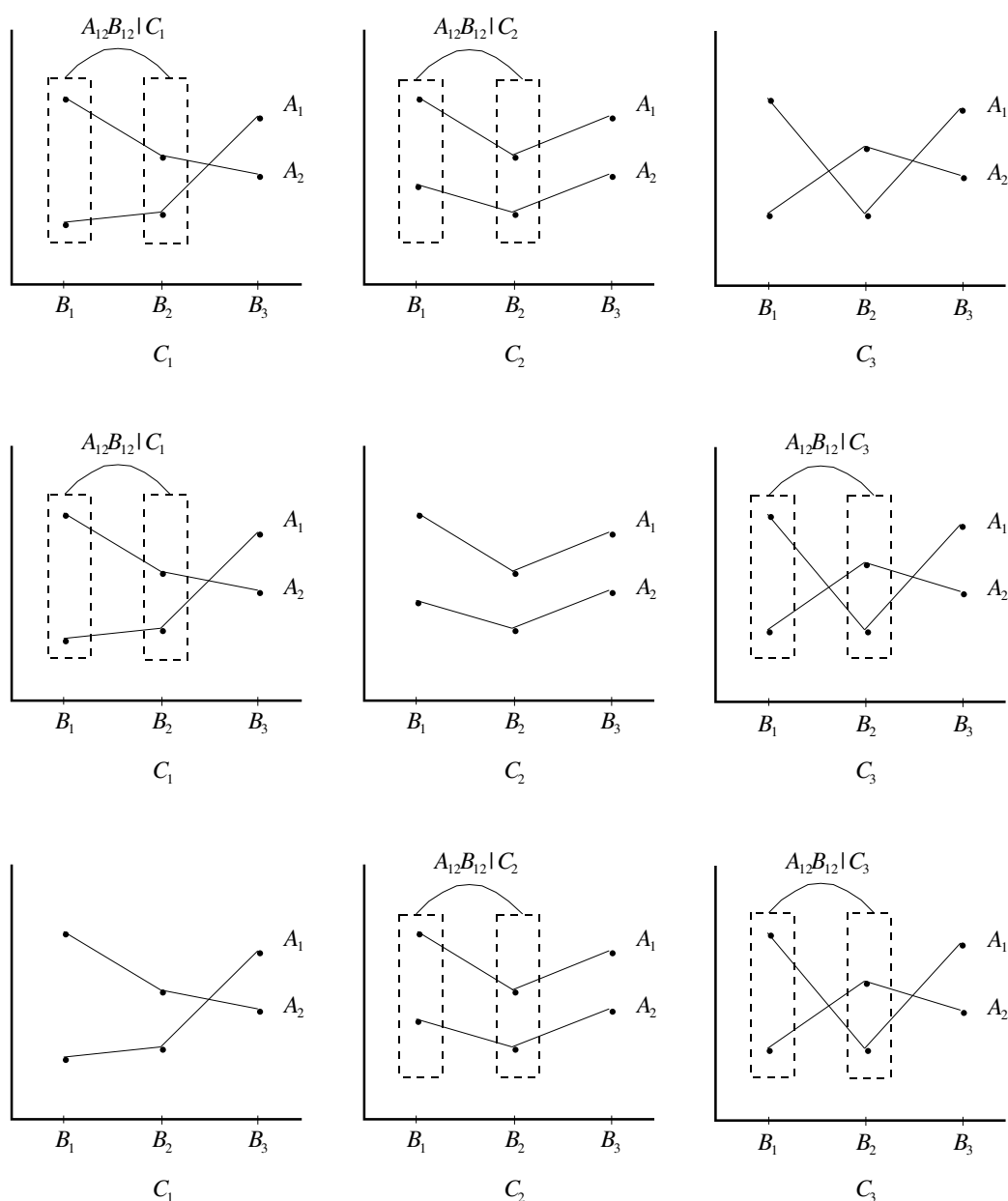
$$\begin{aligned}
A_{12}B_{23}|C_1 &= \mu_{121} - \mu_{221} - (\mu_{131} - \mu_{231}) \\
A_{12}B_{12}|C_2 &= \mu_{112} - \mu_{212} - (\mu_{122} - \mu_{222}) \\
A_{12}B_{13}|C_2 &= \mu_{112} - \mu_{212} - (\mu_{132} - \mu_{232}) \\
A_{12}B_{23}|C_2 &= \mu_{122} - \mu_{222} - (\mu_{132} - \mu_{232}) \\
A_{12}B_{12}|C_3 &= \mu_{113} - \mu_{213} - (\mu_{123} - \mu_{223}) \\
A_{12}B_{13}|C_3 &= \mu_{113} - \mu_{213} - (\mu_{133} - \mu_{233}) \\
A_{12}B_{23}|C_3 &= \mu_{123} - \mu_{223} - (\mu_{133} - \mu_{233})
\end{aligned} \tag{5.9}$$

Con estos 9 efectos simples de interacción es posible hacer 36 comparaciones por pares. Sin embargo, no todas esas comparaciones tienen sentido desde el punto de vista de la interpretación de la interacción de segundo orden. Recordemos una vez más que el objetivo que se está persiguiendo es el de interpretar la interacción de segundo orden. Y recordemos también que una interacción de segundo orden está indicando que los efectos simples de interacción entre cualesquiera dos factores no son iguales en todos los niveles del tercer factor. Por tanto, lo que realmente interesa averiguar es si un efecto simple de interacción concreto es o no el mismo en los diferentes niveles del tercer factor. Interesa averiguar, por ejemplo, si los efectos simples correspondientes a la combinación  $A_{12}B_{12}$  son o no iguales en los tres niveles del factor  $C$ ; pero las comparaciones entre los tres efectos simples  $A_{12}B_{12}$  asociados al mismo nivel del factor  $C$  no tienen utilidad (pues estas comparaciones no están informando acerca de si los efectos simples de la interacción  $AB$  son o no iguales en los diferentes niveles del factor  $C$ ). Los gráficos de la Figura 5.3 pueden ayudar a entender y a concretar cuáles son las comparaciones entre efectos simples de interacción que interesa llevar a cabo.

Cada grupo (fila) de tres gráficos muestra la interacción  $AB$  en cada nivel de  $C$ . Los tres gráficos de la parte superior recogen la comparación entre el efecto simple  $A_{12}B_{12}|C_1$  (las cuatro medias incluidas en los recuadros punteados del primer gráfico) y

el efecto simple  $A_{12}B_{12}|C_2$  (las cuatro medias incluidas en los recuadros punteados del segundo gráfico). Los tres gráficos de la parte media recogen la comparación entre los efectos simples de interacción  $A_{12}B_{12}|C_1$  y  $A_{12}B_{12}|C_3$ . Y los tres gráficos de la parte inferior recogen la comparación entre los efectos simples de interacción  $A_{12}B_{12}|C_2$  y  $A_{12}B_{12}|C_3$ . Lo que se está haciendo con estas tres comparaciones es averiguar si el efecto simple de interacción asociado a la combinación  $A_{12}B_{12}$  es o no el mismo en los tres niveles del factor  $C$ .

**Figura 5.3. Comparaciones entre efectos simples de interacción**



Exactamente lo mismo hay que hacer con los efectos simples de interacción asociados a la combinación  $A_{12}B_{13}$  (otras tres comparaciones) y con los asociados a la combinación  $A_{12}B_{23}$  (otras tres comparaciones) para terminar de comparar entre sí, por pares, los efectos simples de la interacción  $AB$  en cada nivel del factor  $C$ . La Tabla 5.22 ofrece los coeficientes asociados a cada una de estas nueve comparaciones.

Esta estrategia basada en nueve comparaciones permite averiguar si la interacción de primer orden  $AB$  es o no constante (la misma) en todos los niveles del factor  $C$  y, en el caso de que no lo sea, precisar dónde se encuentran las diferencias.

Es importante señalar que no importa qué interacción de primer orden ( $AB$ ,  $AC$ ,  $BC$ ) se tome como referente para la interpretación de la interacción de segundo orden. Con todas ellas se obtiene el mismo resultado. Así, si en lugar de evaluar el comportamiento de la interacción  $AB$  en los diferentes niveles de  $C$ , se prefiere evaluar el comportamiento de la interacción  $AC$  en los diferentes niveles de  $B$ , la situación es idéntica a la descrita. El número de efectos simples de interacción  $AC$  siguen siendo tres en cada nivel de  $B$ :

$$\begin{aligned}A_{12}C_{12}|B_1 &= \mu_{111} - \mu_{211} - (\mu_{112} - \mu_{212}) \\A_{12}C_{13}|B_1 &= \mu_{111} - \mu_{211} - (\mu_{113} - \mu_{213}) \\A_{12}C_{23}|B_1 &= \mu_{112} - \mu_{212} - (\mu_{113} - \mu_{213}) \\A_{12}C_{12}|B_2 &= \mu_{121} - \mu_{221} - (\mu_{122} - \mu_{222}) \\A_{12}C_{13}|B_2 &= \mu_{121} - \mu_{221} - (\mu_{123} - \mu_{223}) \\A_{12}C_{23}|B_2 &= \mu_{122} - \mu_{222} - (\mu_{123} - \mu_{223}) \\A_{12}C_{12}|B_3 &= \mu_{131} - \mu_{231} - (\mu_{132} - \mu_{232}) \\A_{12}C_{13}|B_3 &= \mu_{131} - \mu_{231} - (\mu_{133} - \mu_{233}) \\A_{12}C_{23}|B_3 &= \mu_{132} - \mu_{232} - (\mu_{133} - \mu_{233})\end{aligned}\tag{5.10}$$

**Tabla 5.22. Coeficientes para efectuar comparaciones entre efectos simples de interacción en un diseño 2×3×3**

	$\mu_{111}$	$\mu_{112}$	$\mu_{113}$	$\mu_{121}$	$\mu_{122}$	$\mu_{123}$	$\mu_{131}$	$\mu_{132}$	$\mu_{133}$	$\mu_{211}$	$\mu_{212}$	$\mu_{213}$	$\mu_{221}$	$\mu_{222}$	$\mu_{223}$	$\mu_{231}$	$\mu_{232}$	$\mu_{233}$
$\Psi_1 = A_{12}B_{12} C_1 - A_{12}B_{12} C_2$	1	-1	0	-1	1	0	0	0	0	-1	1	0	1	-1	0	0	0	0
$\Psi_2 = A_{12}B_{12} C_1 - A_{12}B_{12} C_3$	1	0	-1	-1	0	1	0	0	0	-1	0	1	1	0	-1	0	0	0
$\Psi_3 = A_{12}B_{12} C_2 - A_{12}B_{12} C_3$	0	1	-1	0	-1	1	0	0	0	0	-1	1	0	1	-1	0	0	0
$\Psi_4 = A_{12}B_{13} C_1 - A_{12}B_{12} C_2$	1	-1	0	0	0	0	-1	1	0	-1	1	0	0	0	0	1	-1	0
$\Psi_5 = A_{12}B_{13} C_1 - A_{12}B_{12} C_3$	1	0	-1	0	0	0	-1	0	1	-1	0	1	0	0	0	1	0	-1
$\Psi_6 = A_{12}B_{13} C_2 - A_{12}B_{12} C_3$	0	1	-1	0	0	0	0	-1	1	0	-1	1	0	0	0	0	1	-1
$\Psi_7 = A_{12}B_{23} C_1 - A_{12}B_{12} C_2$	0	0	0	1	-1	0	-1	1	0	0	0	0	-1	1	0	1	-1	0
$\Psi_8 = A_{12}B_{23} C_1 - A_{12}B_{12} C_3$	0	0	0	1	0	-1	-1	0	1	0	0	0	-1	0	1	1	0	-1
$\Psi_9 = A_{12}B_{23} C_2 - A_{12}B_{12} C_3$	0	0	0	0	1	-1	0	-1	1	0	0	0	0	-1	1	0	1	-1

$AB|C$

$AC|B$

$BC|A$

$$\Psi_1 = A_{12}B_{12}|C_1 - A_{12}B_{12}|C_2 = A_{12}C_{12}|B_1 - A_{12}C_{12}|B_2 = B_{12}C_{12}|A_1 - B_{12}C_{12}|A_2$$

$$\Psi_2 = A_{12}B_{12}|C_1 - A_{12}B_{12}|C_3 = A_{12}C_{13}|B_1 - A_{12}C_{13}|B_2 = B_{12}C_{13}|A_1 - B_{12}C_{13}|A_2$$

$$\Psi_3 = A_{12}B_{12}|C_2 - A_{12}B_{12}|C_3 = A_{12}C_{23}|B_1 - A_{12}C_{23}|B_2 = B_{12}C_{23}|A_1 - B_{12}C_{23}|A_2$$

$$\Psi_4 = A_{12}B_{13}|C_1 - A_{12}B_{12}|C_2 = A_{12}C_{12}|B_1 - A_{12}C_{12}|B_3 = B_{13}C_{12}|A_1 - B_{13}C_{12}|A_2$$

$$\Psi_5 = A_{12}B_{13}|C_1 - A_{12}B_{12}|C_3 = A_{12}C_{13}|B_1 - A_{12}C_{13}|B_3 = B_{13}C_{13}|A_1 - B_{13}C_{13}|A_2$$

$$\Psi_6 = A_{12}B_{13}|C_2 - A_{12}B_{12}|C_3 = A_{12}C_{23}|B_1 - A_{12}C_{23}|B_3 = B_{13}C_{23}|A_1 - B_{13}C_{23}|A_2$$

$$\Psi_7 = A_{12}B_{23}|C_1 - A_{12}B_{12}|C_2 = A_{12}C_{12}|B_2 - A_{12}C_{12}|B_3 = B_{23}C_{12}|A_1 - B_{23}C_{12}|A_2$$

$$\Psi_8 = A_{12}B_{23}|C_1 - A_{12}B_{12}|C_3 = A_{12}C_{13}|B_2 - A_{12}C_{13}|B_3 = B_{23}C_{13}|A_1 - B_{23}C_{13}|A_2$$

$$\Psi_9 = A_{12}B_{23}|C_2 - A_{12}B_{12}|C_3 = A_{12}C_{23}|B_2 - A_{12}C_{23}|B_3 = B_{23}C_{23}|A_1 - B_{23}C_{23}|A_2$$

Y las comparaciones que es necesario efectuar para comparar por pares estos nueve efectos siguen siendo las mismas nueve comparaciones propuestas en la Tabla 5.22 (la mitad inferior de la tabla muestra la equivalencia entre estas comparaciones y las correspondientes a la combinación  $AB|C$ ).

Por último, si en lugar de la interacción  $AB$  o la  $AC$ , se prefiere evaluar el comportamiento de la interacción  $BC$  en los diferentes niveles del factor  $A$ , la situación tampoco cambia. Aunque ahora el número de efectos simples de interacción  $BC$  son nueve en cada nivel de  $A$  (18 efectos en total):

$$\begin{aligned}
 B_{12}C_{12}|A_1 &= \mu_{111} - \mu_{121} - (\mu_{112} - \mu_{122}) \\
 B_{12}C_{13}|A_1 &= \mu_{111} - \mu_{121} - (\mu_{113} - \mu_{123}) \\
 B_{12}C_{23}|A_1 &= \mu_{112} - \mu_{122} - (\mu_{113} - \mu_{123}) \\
 B_{13}C_{12}|A_1 &= \mu_{111} - \mu_{131} - (\mu_{112} - \mu_{132}) \\
 B_{13}C_{13}|A_1 &= \mu_{111} - \mu_{131} - (\mu_{113} - \mu_{133}) \\
 B_{13}C_{23}|A_1 &= \mu_{112} - \mu_{132} - (\mu_{113} - \mu_{133}) \\
 B_{23}C_{12}|A_1 &= \mu_{121} - \mu_{131} - (\mu_{122} - \mu_{132}) \\
 B_{23}C_{13}|A_1 &= \mu_{121} - \mu_{131} - (\mu_{123} - \mu_{133}) \\
 B_{23}C_{23}|A_1 &= \mu_{122} - \mu_{132} - (\mu_{123} - \mu_{133}) \\
 \\ 
 B_{12}C_{12}|A_2 &= \mu_{211} - \mu_{221} - (\mu_{212} - \mu_{222}) \\
 B_{12}C_{13}|A_2 &= \mu_{211} - \mu_{221} - (\mu_{213} - \mu_{223}) \\
 B_{12}C_{23}|A_2 &= \mu_{212} - \mu_{222} - (\mu_{213} - \mu_{223}) \\
 B_{13}C_{12}|A_2 &= \mu_{211} - \mu_{231} - (\mu_{212} - \mu_{232}) \\
 B_{13}C_{13}|A_2 &= \mu_{211} - \mu_{231} - (\mu_{213} - \mu_{233}) \\
 B_{13}C_{23}|A_2 &= \mu_{212} - \mu_{232} - (\mu_{213} - \mu_{233}) \\
 B_{23}C_{12}|A_2 &= \mu_{221} - \mu_{231} - (\mu_{222} - \mu_{232}) \\
 B_{23}C_{13}|A_2 &= \mu_{221} - \mu_{231} - (\mu_{223} - \mu_{233}) \\
 B_{23}C_{23}|A_2 &= \mu_{222} - \mu_{232} - (\mu_{223} - \mu_{233}),
 \end{aligned} \tag{5.11}$$

y las comparaciones que hay que diseñar para compararlos por pares siguen siendo las mismas nueve ya propuestas en la Tabla 5.22: hay que comparar cada efecto en  $A_1$  con su correspondiente efecto en  $A_2$  (la mitad inferior de la Tabla 5.22 muestra la equivalencia entre estas comparaciones y las correspondientes a las combinaciones  $AB|C$  y  $AC|B$ ).

### 5.2.3. Diseños $3 \times 3 \times 3$

La Tabla 5.23 muestra la notación utilizada en un diseño  $3 \times 3 \times 3$ . En el apartado anterior se ha tenido ocasión de comprobar que, al añadir un solo nivel a uno de los factores, aumenta rápidamente el número de comparaciones que es posible plantear.

**Tabla 5.23. Notación utilizada en un diseño  $3 \times 3 \times 3$**

	$B_1$			$B_2$			$B_3$		
	$C_1$	$C_2$	$C_3$	$C_1$	$C_2$	$C_3$	$C_1$	$C_2$	$C_3$
$A_1$	$\mu_{111}$	$\mu_{112}$	$\mu_{113}$	$\mu_{121}$	$\mu_{122}$	$\mu_{123}$	$\mu_{131}$	$\mu_{132}$	$\mu_{133}$
$A_2$	$\mu_{211}$	$\mu_{212}$	$\mu_{213}$	$\mu_{221}$	$\mu_{222}$	$\mu_{223}$	$\mu_{231}$	$\mu_{232}$	$\mu_{233}$
$A_3$	$\mu_{311}$	$\mu_{312}$	$\mu_{313}$	$\mu_{321}$	$\mu_{322}$	$\mu_{323}$	$\mu_{331}$	$\mu_{332}$	$\mu_{333}$

En un diseño  $2 \times 3 \times 3$  como el estudiado en el apartado anterior, los efectos simples asociados a la interacción  $AB$  son 3 en cada nivel del factor  $C$ . Esto arroja un total de 9 efectos simples de interacción (ver ecuación [5.9]).

En un diseño  $3 \times 3 \times 3$ , el número de efectos simples asociados a la interacción  $AB$  son 9 en cada nivel del factor  $C$ . Así, en el nivel  $C_1$ , por ejemplo, los 9 efectos simples de interacción asociados a la combinación  $AB$  son los siguientes:

$$A_{12}B_{12}|C_1 = \mu_{111} - \mu_{211} - (\mu_{121} - \mu_{221})$$

$$A_{12}B_{13}|C_1 = \mu_{111} - \mu_{211} - (\mu_{131} - \mu_{231})$$



$$\begin{aligned}A_{12}B_{23}|C_1 &= \mu_{121} - \mu_{221} - (\mu_{131} - \mu_{231}) \\A_{13}B_{12}|C_1 &= \mu_{111} - \mu_{311} - (\mu_{121} - \mu_{321}) \\A_{13}B_{13}|C_1 &= \mu_{111} - \mu_{311} - (\mu_{131} - \mu_{331}) \\A_{13}B_{23}|C_1 &= \mu_{121} - \mu_{321} - (\mu_{131} - \mu_{331}) \\A_{23}B_{12}|C_1 &= \mu_{211} - \mu_{311} - (\mu_{221} - \mu_{321}) \\A_{23}B_{13}|C_1 &= \mu_{211} - \mu_{311} - (\mu_{231} - \mu_{331}) \\A_{23}B_{23}|C_1 &= \mu_{221} - \mu_{321} - (\mu_{231} - \mu_{331})\end{aligned}\tag{5.12}$$

En el nivel  $C_2$  hay otros 9 efectos simples de interacción  $AB$  (que se obtienen cambiando el último subíndice de todas las medias por un 2); y otros 9 en el nivel  $C_3$  (que se obtienen cambiando el último subíndice de todas las medias por un 3). Es decir, en un diseño  $3 \times 3 \times 3$  hay un total de 27 efectos simples de interacción.

Con estos 27 efectos es posible diseñar 351 comparaciones por pares. Pero, según se ha argumentado ya, no todas estas comparaciones tienen sentido. Siguiendo la lógica expuesta en el apartado anterior, lo que interesa averiguar es si un efecto simple de interacción concreto es o no el mismo en los diferentes niveles del tercer factor. Interesa averiguar, por ejemplo, si los efectos simples correspondientes a la combinación  $A_{12}B_{12}$  son o no iguales en los tres niveles del factor  $C$ ; o si los efectos simples correspondientes a la combinación  $A_{12}B_{13}$  son o no iguales en los tres niveles del factor  $C$ ; etc. Pero las comparaciones entre los 9 efectos simples asociados al mismo nivel del factor  $C$  no tienen utilidad, pues esas comparaciones no están informando acerca de cómo los efectos simples de la interacción  $AB$  se comportan en los diferentes niveles de  $C$ .

En consecuencia, interesa averiguar si cada uno de los 9 efectos simples asociados a la interacción  $AB$  (los definidos en [5.12]) es o no el mismo en los diferentes niveles del factor  $C$  (por supuesto, daría exactamente igual preguntarse por la interacción  $AC$  en los diferentes niveles del factor  $B$ , o por la interacción  $BC$  en los diferentes niveles del

factor A). Por tanto, las comparaciones que interesa efectuar son las que permiten evaluar el comportamiento de cada uno de los 9 efectos simples de interacción  $AB$  al comparar lo que ocurre con ellos en los niveles  $C_1$  y  $C_2$ , en los niveles  $C_1$  y  $C_3$ , y en los niveles  $C_2$  y  $C_3$ . Es decir, para poder interpretar correctamente el significado de una interacción de segundo orden en un diseño  $3 \times 3 \times 3$ , es necesario llevar a cabo un total de 27 comparaciones: 3 por cada uno de los 9 efectos simples de interacción  $AB$  (o  $AC$ , o  $BC$ ). La Tabla 5.24 ofrece los coeficientes asociados a cada una de estas comparaciones.

Por supuesto, dado que no todas estas comparaciones son independientes entre sí, la presencia de hipótesis orientadas hacia determinados efectos puede ayudar a reducir sensiblemente el número de comparaciones que es necesario llevar a cabo. En este sentido, debe tenerse siempre en cuenta que la presencia de una interacción de segundo orden está indicando que ninguna de las tres interacciones de primer orden es constante (la misma) en todos los niveles del tercer factor. Consecuentemente, la atención del investigador debe ir dirigida a la comprobación de esa circunstancia. Y las hipótesis específicas que pueden contribuir a reducir el número de comparaciones son las hipótesis referidas al comportamiento de *alguna* interacción de primer orden en *alguno* de los niveles del tercer factor.

**Tabla 5.24. Coeficientes para efectuar comparaciones entre efectos simples de interacción en un diseño 3×3×3**

	$\mu_{11}$	$\mu_{112}$	$\mu_{113}$	$\mu_{121}$	$\mu_{122}$	$\mu_{123}$	$\mu_{131}$	$\mu_{132}$	$\mu_{133}$	$\mu_{211}$	$\mu_{212}$	$\mu_{213}$	$\mu_{221}$	$\mu_{222}$	$\mu_{223}$	$\mu_{231}$	$\mu_{232}$	$\mu_{233}$	$\mu_{311}$	$\mu_{312}$	$\mu_{313}$	$\mu_{321}$	$\mu_{322}$	$\mu_{323}$	$\mu_{331}$	$\mu_{332}$	$\mu_{333}$
$A_{12}B_{12}C_{12}$	1	-1	0	-1	1	0	0	0	0	-1	1	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0
$A_{12}B_{12}C_{13}$	1	0	-1	-1	0	1	0	0	0	-1	0	1	1	0	-1	0	0	0	0	0	0	0	0	0	0	0	0
$A_{12}B_{12}C_{23}$	0	1	-1	0	-1	1	0	0	0	0	-1	1	0	1	-1	0	0	0	0	0	0	0	0	0	0	0	0
$A_{12}B_{13}C_{12}$	1	-1	0	0	0	0	-1	1	0	-1	1	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0	0
$A_{12}B_{13}C_{13}$	1	0	-1	0	0	0	-1	0	1	-1	0	1	0	0	0	1	0	-1	0	0	0	0	0	0	0	0	0
$A_{12}B_{13}C_{23}$	0	1	-1	0	0	0	0	-1	1	0	-1	1	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	0
$A_{12}B_{23}C_{12}$	0	0	0	1	-1	0	-1	1	0	0	0	0	-1	1	0	1	-1	0	0	0	0	0	0	0	0	0	0
$A_{12}B_{23}C_{13}$	0	0	0	1	0	-1	-1	0	1	0	0	0	-1	0	1	1	0	-1	0	0	0	0	0	0	0	0	0
$A_{12}B_{23}C_{23}$	0	0	0	0	1	-1	0	-1	1	0	0	0	0	-1	1	0	1	-1	0	0	0	0	0	0	0	0	0
$A_{13}B_{12}C_{12}$	1	-1	0	-1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	1	0	1	-1	0	0	0	0
$A_{13}B_{12}C_{13}$	1	0	-1	-1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	1	1	0	-1	0	0	0
$A_{13}B_{12}C_{23}$	0	1	-1	0	-1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	1	0	1	-1	0	0	0
$A_{13}B_{13}C_{12}$	1	-1	0	0	0	0	-1	1	0	0	0	0	0	0	0	0	0	0	-1	1	0	0	0	0	1	-1	0
$A_{13}B_{13}C_{13}$	1	0	-1	0	0	0	-1	0	1	0	0	0	0	0	0	0	0	0	-1	0	1	0	0	0	1	0	-1
$A_{13}B_{13}C_{23}$	0	1	-1	0	0	0	0	-1	1	0	0	0	0	0	0	0	0	0	0	-1	1	0	0	0	0	1	-1
$A_{13}B_{23}C_{12}$	0	0	0	1	-1	0	-1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	1	0	1	-1	0
$A_{13}B_{23}C_{13}$	0	0	0	1	0	-1	-1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	1	1	0	-1
$A_{13}B_{23}C_{23}$	0	0	0	0	1	-1	0	-1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	1	0	1	-1
$A_{23}B_{12}C_{12}$	0	0	0	0	0	0	0	0	0	1	-1	0	-1	1	0	0	0	0	-1	1	0	1	-1	0	0	0	0
$A_{23}B_{12}C_{13}$	0	0	0	0	0	0	0	0	0	1	0	-1	-1	0	1	0	0	0	-1	0	1	1	0	-1	0	0	0
$A_{23}B_{12}C_{23}$	0	0	0	0	0	0	0	0	0	0	1	-1	0	-1	1	0	0	0	0	-1	1	0	1	-1	0	0	0
$A_{23}B_{13}C_{12}$	0	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	-1	1	0	-1	1	0	0	0	0	1	-1	0
$A_{23}B_{13}C_{13}$	0	0	0	0	0	0	0	0	0	1	0	-1	0	0	0	-1	0	1	-1	0	1	0	0	0	1	0	-1
$A_{23}B_{13}C_{23}$	0	0	0	0	0	0	0	0	0	0	1	-1	0	0	0	0	-1	1	0	-1	1	0	0	0	0	1	-1
$A_{23}B_{23}C_{12}$	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	0	-1	1	0	0	0	0	-1	1	0	1	-1	0
$A_{23}B_{23}C_{13}$	0	0	0	0	0	0	0	0	0	0	0	0	1	0	-1	-1	0	1	0	0	0	-1	0	1	1	0	-1
$A_{23}B_{23}C_{23}$	0	0	0	0	0	0	0	0	0	0	0	0	0	1	-1	0	-1	1	0	0	0	0	-1	1	0	1	-1

Nota:  $A_{12}B_{12}C_{12}$  representa, indistintamente, cualquiera de estas tres comparaciones: (1)  $A_{12}B_{12}|C_1 - A_{12}B_{12}|C_2$ ; (2)  $A_{12}C_{12}|B_1 - A_{12}C_{12}|B_2$ ; y (3)  $B_{12}C_{12}|A_1 - B_{12}C_{12}|A_2$ . Del mismo modo,  $A_{12}B_{12}C_{13}$  representa, indistintamente, cualquiera de estas tres comparaciones: (1)  $A_{12}B_{12}|C_1 - A_{12}B_{12}|C_3$ ; (2)  $A_{12}C_{13}|B_1 - A_{12}C_{13}|B_2$ ; y (3)  $B_{12}C_{13}|A_1 - B_{12}C_{13}|A_2$ . Etc.

#### 5.2.4. El procedimiento ANOVA de un factor

La utilidad del tipo de comparaciones propuestas en los apartados anteriores radica en que están informando acerca de si una interacción de primer orden es o no constante (la misma) en cada nivel del tercer factor (pues cada comparación involucra ocho medias). Pero, además, al tratarse de comparaciones que tienen asociado un único grado de libertad, poseen la importante ventaja de que cualquier hipótesis nula del tipo  $\psi_h = 0$  puede ponerse a prueba utilizando el estadístico  $t$  convencional:

$$t = \frac{\hat{\Psi}_h}{\sqrt{MCE \sum_k \frac{c_k^2}{n_k}}}$$

aplicando la correspondiente corrección para controlar la tasa de error asociada al número de comparaciones.

Este estadístico, idéntico al propuesto en el apartado anterior para las comparaciones referidas a la interacción de primer orden, se distribuye según el modelo de probabilidad  $t$  de *Student* con grados de libertad igual al número de casos menos el número de casillas ( $h$  se refiere a una comparación cualquiera;  $k$  se refiere al número de casillas del diseño; y  $MCE$  es la media cuadrática error o error cuadrático medio). Sirve aquí lo dicho en el apartado anterior sobre la posibilidad de aplicar este estadístico con el SPSS.

Para ilustrar cómo utilizar el programa estadístico SPSS para efectuar algunas de las comparaciones propuestas se desarrolla un ejemplo práctico basado en un diseño  $2 \times 2 \times 3$ .

Para estudiar el posible efecto que tiene sobre la agresividad una determinada sustancia psicoactiva generalmente consumida para uso recreativo, se ha diseñado un

experimento en el que cada unidad experimental está formada por un par de ratones machos. A un animal de cada pareja se le administra una dosis de la sustancia psicoactiva (grupo experimental) o una dosis de suero fisiológico (grupo control). Durante las dos horas posteriores a la administración de las sustancias de prueba se ha registrando el número de episodios de agresividad observados entre cada pareja de ratones. Se ha manipulado el hacinamiento utilizando jaulas de dos tamaños (grandes y pequeñas). Finalmente, puesto que los ratones son animales de hábitos nocturnos, con un grupo de ratones se ha realizado el experimento por la mañana, con otro por la tarde y con otro más por la noche. La expectativa del investigador era que se observarían más episodios de agresividad en el grupo experimental que en el control y que el número de episodios agresivos sería mayor en las jaulas pequeñas que en las grandes, pudiendo aumentar ambos efectos durante el periodo de mayor actividad, es decir, por la noche.

Se trata de un diseño de tres factores  $2 \times 2 \times 3$  (*tratamiento*: experimental, control; *hacinamiento*: bajo, alto; y *momento*: mañana, tarde, y noche) de efectos fijos, completamente al azar. Se han asignado 5 pares de ratones a cada una de las condiciones experimentales. La Tabla 5.25 muestra el número de episodios de agresividad observado en cada unidad experimental (la tabla también incluye, en la columna *factor*, los valores asignados a cada combinación de niveles de los tres factores para poder efectuar con el SPSS las comparaciones lineales propuestas en los apartados anteriores).

**Tabla 5.25. Datos del ejemplo**

Tto	Hacin.	Mom.	Episodios	Factor	Tto	Hacin.	Mom.	Episodios	Factor
1	1	1	0	1	2	1	1	0	7
1	1	1	4	1	2	1	1	2	7
1	1	1	5	1	2	1	1	5	7
1	1	1	3	1	2	1	1	4	7
1	1	1	4	1	2	1	1	1	7
1	1	2	4	2	2	1	2	2	8
1	1	2	7	2	2	1	2	5	8
1	1	2	8	2	2	1	2	6	8
1	1	2	2	2	2	1	2	3	8
1	1	2	4	2	2	1	2	7	8
1	1	3	13	3	2	1	3	6	9
1	1	3	10	3	2	1	3	8	9
1	1	3	8	3	2	1	3	9	9
1	1	3	14	3	2	1	3	9	9
1	1	3	12	3	2	1	3	12	9
1	2	1	17	4	2	2	1	1	10
1	2	1	22	4	2	2	1	3	10
1	2	1	21	4	2	2	1	4	10
1	2	1	19	4	2	2	1	3	10
1	2	1	19	4	2	2	1	6	10
1	2	2	22	5	2	2	2	6	11
1	2	2	20	5	2	2	2	1	11
1	2	2	23	5	2	2	2	5	11
1	2	2	18	5	2	2	2	5	11
1	2	2	17	5	2	2	2	7	11
1	2	3	21	6	2	2	3	10	12
1	2	3	17	6	2	2	3	8	12
1	2	3	23	6	2	2	3	9	12
1	2	3	22	6	2	2	3	14	12
1	2	3	20	6	2	2	3	12	12

*Tto* (tratamiento): 1 = experimental, 2 = control.

*Hacin.* (grado de hacinamiento): 1 = bajo, 2 = alto.

*Mom.*(momento del día): 1 = mañana, 2 = tarde, 3 = noche.

*Episodios*: número de episodios agresivos.

*Factor*: códigos únicos para cada combinación entre *tratamiento*, *hacinamiento* y *momento*.

La Tabla 5.26 muestra las medias obtenidas en cada combinación de niveles. Y la Tabla 5.27 recoge los resultados del ANOVA, los cuales revelan, entre otras cosas, un efecto significativo de la interacción de segundo orden ( $p = 0,01$ ).

Tabla 5.26. Medias observadas en cada combinación de niveles

	Hacinamiento bajo			Hacinamiento alto		
	Mañana	Tarde	Noche	Mañana	Tarde	Noche
Experim.	3,20 (1,92)	5,00 (2,45)	11,40 (2,41)	19,60 (1,95)	20,00 (2,55)	20,60 (2,30)
Control	2,40 (2,07)	4,60 (2,07)	8,80 (2,17)	3,40 (1,82)	4,80 (2,28)	10,60 (2,41)

Tabla 5.27. Tabla resumen del ANOVA

<i>Fuente de variación</i>	<i>SC tipo III</i>	<i>gl</i>	<i>MC</i>	<i>F</i>	<i>p</i>
Tto	851,267	1	851,267	174,024	0,000
Hacinamiento	792,067	1	792,067	161,922	0,000
Momento	351,033	2	175,517	35,881	0,000
Tto*Hacinamiento	589,067	1	589,067	120,422	0,000
Tto*Momento	12,633	2	6,317	1,291	0,284
Hacinamiento*Momento	26,433	2	13,217	2,702	0,077
Tto*Hacinamiento*Momento	49,633	2	24,817	5,073	0,010
Error	234,800	48	4,892		
Total	2.906,933	59			

El hecho de que la interacción de segundo orden sea significativa está indicando que ninguna de las interacciones de primer orden es constante (no es la misma) en todos los niveles del tercer factor. Es decir, está indicando que la interacción entre *tratamiento* y *hacinamiento* no es la misma en los tres *momentos*, que la interacción entre *tratamiento* y *momento* no es la misma en los dos niveles de *hacinamiento*, y que la interacción entre *hacinamiento* y *momento* no es la misma en los dos *tratamientos*.

Ahora bien, para interpretar la interacción de segundo orden puede tomarse como referente cualquiera de las interacciones de primer orden pues, según se ha señalado ya, las comparaciones que es necesario llevar a cabo son exactamente las mismas cualquiera que sea la interacción de primer orden que se tome como referente. Puesto que los factores *tratamiento* y *hacinamiento* tienen sólo dos niveles, comenzaremos tomando como referente la interacción entre esos dos factores. Esto significa que debemos com-

parar entre sí, por pares, los efectos simples de interacción asociados a la combinación *tratamiento-hacinamiento*, lo cual requiere llevar a cabo las tres comparaciones propuestas en la ecuación [5.6]. Los coeficientes correspondientes a estas tres comparaciones están recogidos en la Tabla 5.20 (en  $\psi_1$ ,  $\psi_2$  y  $\psi_3$ ; para que la correspondencia con las comparaciones propuestas en la Tabla 5.20 sea correcta, estamos llamando *A* al factor *tratamiento*, *B* al factor *hacinamiento* y *C* al factor *momento*).

Para poder realizar estas comparaciones con el SPSS hay que comenzar creando una variable cuyos códigos representen las 12 condiciones experimentales del diseño: un 1 para la combinación  $A_1B_1C_1$ , un 2 para la combinación  $A_1B_1C_2$ , ..., un 12 para la combinación  $A_2B_2C_3$ . La Tabla 5.25 muestra estos códigos en la columna *factor*. Utilizando esta nueva variable como factor en el procedimiento SPSS *ANOVA de un factor* y asignando en el cuadro de diálogo *Contrastes* los coeficientes correspondientes a las comparaciones  $\psi_1$ ,  $\psi_2$  y  $\psi_3$  (ver Tabla 5.20), se obtienen los resultados que muestra la Tabla 5.28 (como se están realizando tres comparaciones, el nivel de significación por comparación puede establecerse en  $0,05/3 = 0,017$ ).

Los resultados muestran dos comparaciones significativas ( $\psi_2$  y  $\psi_3$ ): la interacción entre *tratamiento* y *hacinamiento* en el momento 3 = *noche* no es la misma que en los momentos 1 = *mañana* ( $p = 0,006$ ) y 2 = *tarde* ( $p = 0,011$ ). La primera comparación ( $\psi_1$ ) no es significativa ( $p = 0,813$ ); por tanto, no puede afirmarse que la interacción entre *tratamiento* y *hacinamiento* sea distinta en los momentos 1 = *mañana* y 2 = *tarde*.

**Tabla 5.28. Comparaciones entre los efectos simples de interacción**

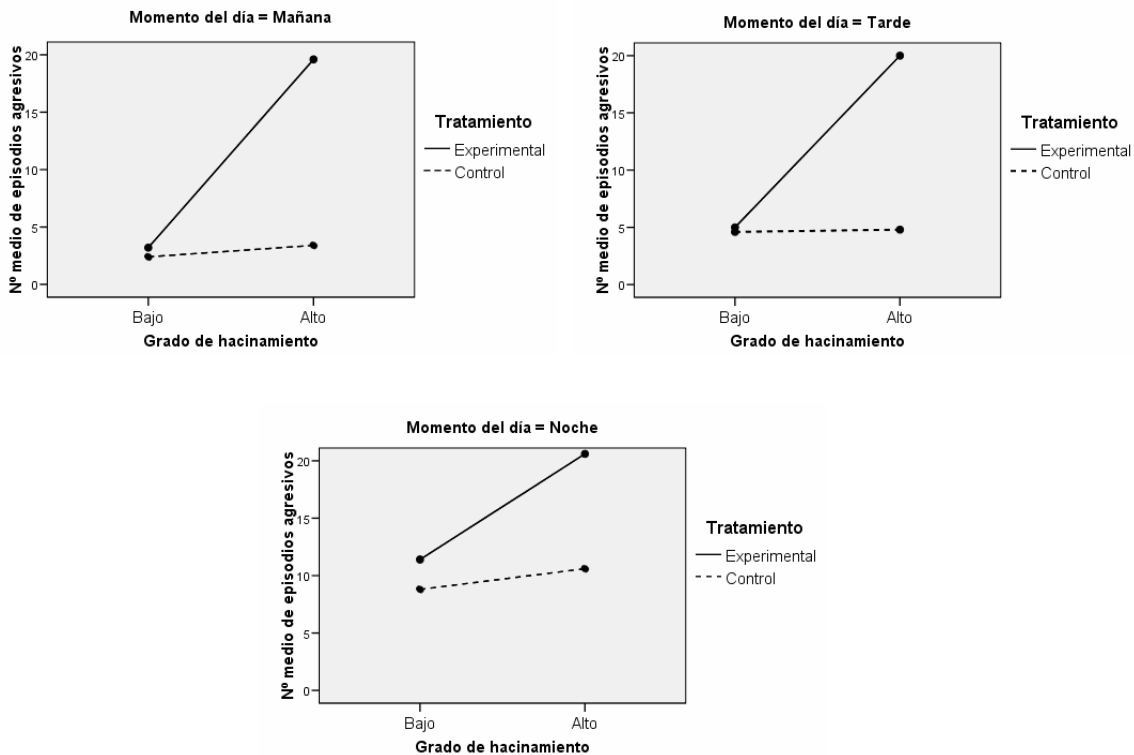
	$\hat{\psi}_k$	Err. típ.	<i>t</i>	<i>gl</i>	<i>p</i>
$\psi_1 = TH M_1 - TH M_2$	-0,60	2,798	-0,214	48	0,831
$\psi_2 = TH M_1 - TH M_3$	-8,00	2,798	-2,860	48	0,006
$\psi_3 = TH M_2 - TH M_3$	-7,40	2,798	-2,645	48	0,011

*T* = tratamiento, *H* = hacinamiento, *M* = momento.



Para ayudar a entender lo que está ocurriendo, los gráficos de líneas de la Figura 5.4 representan la interacción entre *tratamiento* y *hacinamiento* en los tres momentos del día. Los resultados de la Tabla 5.28 indican que la interacción representada en el tercer gráfico (noche) difiere de las representadas en los dos primeros (mañana y tarde). Llamando *efecto del tratamiento* a la diferencia entre los grupos experimental y control, cabe concluir que la diferencia encontrada en el efecto del tratamiento entre las condiciones de hacinamiento bajo y alto por la noche es distinta (en concreto, menor) de la diferencia encontrada por la mañana y por la tarde.

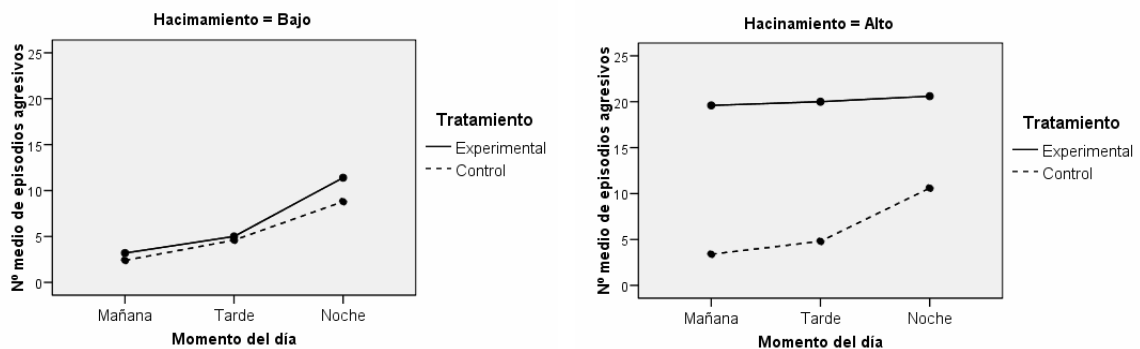
**Figura 5.4** Interacción entre *tratamiento* y *hacinamiento* en cada *momento del día*



Si en lugar de interpretar la interacción de segundo orden tomando como referente la interacción de primer orden entre *tratamiento* y *hacinamiento* se prefiere utilizar otra

interacción de primer orden, las comparaciones que hay que llevar a cabo son las mismas (ver Tabla 5.28) y la conclusión a la que se llega también es la misma. Veamos qué ocurre si se utiliza como referente para la interpretación la interacción entre *tratamiento* y *momento*. La Figura 5.5 muestra los gráficos de líneas correspondientes a esa interacción en cada nivel de *hacinamiento*. El hecho de que la interacción de segundo orden sea significativa está indicando que la interacción representada en el primer gráfico (hacinamiento *bajo*) difiere de la representada en el segundo (hacinamiento *alto*). Llamando *efecto del tratamiento* a la diferencia entre los grupos experimental y control y teniendo en cuenta los resultados de las comparaciones entre efectos simples de interacción (ver Tabla 5.28), cabe concluir, que, por la noche, la diferencia en el efecto del tratamiento entre las condiciones de bajo y alto hacinamiento es distinta (en concreto, menor) de esa misma diferencia por la mañana y por la tarde (conclusión idéntica a la que se ha llegado antes tomando como referente la interacción entre *tratamiento* y *hacinamiento*).

**Figura 5.5. Interacción entre *tratamiento* y *momento* en cada nivel de *hacinamiento***



Aunque el significado de la interacción de segundo orden se agota aquí, la forma habitual de proceder de los investigadores consiste en realizar otro tipo de contrastes que, no sólo no informan sobre la interacción de segundo orden, sino que, al perderla de vista

(normalmente por buscar qué ocurre con otros efectos esperados), se llega a conclusiones no del todo correctas.

Una de estas formas de proceder consiste en evaluar los efectos simples de interacción por separado (estrategia equivalente a evaluar por separado los efectos simples en un diseño de dos factores). Esta estrategia admite dos enfoques analíticos diferentes. El primero de ellos consiste en realizar comparaciones lineales de un grado de libertad (tal como se ha hecho para obtener los resultados de la Tabla 5.28) utilizando los coeficientes que muestra la Tabla 5.29. Con esta forma de proceder (que utiliza en todas las comparaciones la misma media cuadrática que el modelo de ANOVA que incluye todos los efectos posibles del diseño) se obtienen los resultados que muestra la Tabla 5.30.

**Tabla 5.29. Coeficientes para valorar la significación individual de los efectos simples asociados a la interacción entre *tratamiento* y *hacinamiento***

	$\mu_{111}$	$\mu_{112}$	$\mu_{113}$	$\mu_{121}$	$\mu_{122}$	$\mu_{123}$	$\mu_{211}$	$\mu_{212}$	$\mu_{213}$	$\mu_{221}$	$\mu_{222}$	$\mu_{223}$
$TH M_1$	1	0	0	-1	0	0	-1	0	0	1	0	0
$TH M_2$	0	1	0	0	-1	0	0	-1	0	0	1	0
$TH M_3$	0	0	1	0	0	-1	0	0	-1	0	0	1

$T$  = tratamiento,  $H$  = hacinamiento,  $M$  = momento.

**Tabla 5.30. Significación individual de los efectos simples de interacción**

	$\hat{\psi}_k$	<i>Err. típ.</i>	$t$	$gl$	$p$
$TH M_1$	-15,40	1,978	-7,785	48	0,000
$TH M_2$	-14,80	1,978	-7,481	48	0,000
$TH M_3$	-7,40	1,978	-3,560	48	0,000

$T$  = tratamiento,  $H$  = hacinamiento,  $M$  = momento.

Otra forma de evaluar la significación individual de los efectos simples de interacción consiste en llevar a cabo un ANOVA  $2 \times 2$  (*tratamiento* y *hacinamiento*) en cada nivel del tercer factor (*momento*) centrando la atención en la significación de la interacción de primer orden *tratamiento*  $\times$  *hacinamiento*. Con este otro enfoque analítico se está utili-

zando una media cuadrática distinta para valorar cada interacción (es decir, para valorar cada efecto simple de interacción); sin embargo, si se cumple el supuesto de igualdad de varianzas, los resultados que se obtienen son los mismos que con la estrategia basada en comparaciones lineales de un grado de libertad. La Tabla 5.31 muestra los resultados obtenidos con este enfoque.

**Tabla 5.31. Significación de la interacción entre *tratamiento* y *hacinamiento* en cada momento del día (significación individual de los efectos simples de interacción)**

		<i>gl</i>	<i>MC</i>	<i>F</i>	<i>p</i>
Mañana	Tratamiento × Hacinamiento	1	296,450	78,530	0,000
	Error	16	3,775		
Tarde	tratamiento × Hacinamiento	1	273,800	49,782	0,000
	Error	16	5,500		
Noche	tratamiento × Hacinamiento	1	68,450	12,676	0,003
	Error	16	5,400		

Los resultados de las Tablas 5.30 y 5.31 coinciden en señalar que los tres efectos simples asociados a la interacción entre *tratamiento* y *hacinamiento* son significativos. Esto significa que el efecto del tratamiento (la diferencia entre el grupo experimental y el control) es mayor en la condición de hacinamiento alto que en la de hacinamiento bajo; y esto, en los tres momentos considerados (ver Figura 5.4). Puede que esta conclusión tenga algún interés, pero en ella se está perdiendo de vista que la interacción de segundo orden es significativa: aunque el efecto del tratamiento es mayor en la condición de hacinamiento alto que en la de hacinamiento bajo, esta diferencia entre hacinamientos no es la misma en los tres momentos (ver Tabla 5.28). Y el hecho de que los tres efectos simples de interacción sean significativos (Tablas 5.30 y 5.31) no sólo

no aclara nada acerca de la interacción de segundo orden, sino que podría malinterpretarse como que no existe tal interacción.

Recordemos ahora que la expectativa del investigador era, en primer lugar, que se observarían más episodios de agresividad en el grupo experimental que en el control. Esto es cierto si sólo se presta atención al efecto principal del tratamiento ( $p < 0,0005$ ; ver Tabla 5.27), pero no si se presta atención a la información que ofrece la interacción de segundo orden. Al considerar esta información se observa que, aunque en la condición de alto hacinamiento el efecto del tratamiento es significativo en los tres momentos del día\* ( $p < 0,0005$  en los tres momentos), en la condición de bajo hacinamiento no puede afirmarse que exista efecto del tratamiento en ninguno de los tres momentos (mañana:  $p = 0,570$ ; tarde:  $p = 0,776$ ; noche:  $p = 0,069$ ).

La segunda expectativa del investigador era que el número de episodios agresivos sería mayor en las jaulas pequeñas (hacinamiento alto) que en las grandes (hacinamiento bajo). También esto es cierto si únicamente se tiene en cuenta el efecto principal del hacinamiento ( $p < 0,0005$ ; ver Tabla 5.27). Pero no debe olvidarse que la interacción de segundo orden es significativa, lo cual revela que, aunque el número de episodios agresivos es efectivamente mayor en condiciones de hacinamiento alto, esto sólo ocurre en el en el grupo experimental ( $p < 0,0005$  en los tres momentos); en el grupo control, la diferencia entre las condiciones de bajo y alto hacinamiento no es significativa en ninguno de los tres momentos (mañana:  $p = 0,478$ ; tarde:  $p = 0,887$ ; noche:  $p = 0,204$ ).

La última expectativa del investigador se refería al efecto mediador del momento del día. En concreto, esperaba que tanto el efecto del tratamiento como el del hacina-

---

\* Estos resultados pueden obtenerse de diferentes maneras: mediante comparaciones planeadas modificando la sintaxis SPSS del procedimiento UNIANOVA (ver Pardo y Ruiz, 2005, págs. 374-375); con la sentencia LMATRIX (ver Pardo y Ruiz, 2005, págs. 380-385); y diseñando comparaciones de un grado de libertad, como las que venimos realizando en este capítulo, en las que únicamente estén involucradas (con coeficientes 1 y -1) las dos medias que interesa comparar.

miento aumentarían durante el periodo de mayor actividad, es decir, por la noche. Esta expectativa hace referencia al efecto de las interacciones de primer orden entre *tratamiento* y *momento* y entre *hacinamiento* y *momento*. Los resultados del ANOVA (ver Tabla 5.27) indican que ninguna de estas dos interacciones es significativa ( $p = 0,284$  y  $p = 0,077$ , respectivamente). Esto podría llevar a muchos investigadores a pensar que tanto el efecto del tratamiento como el del hacinamiento son independientes del momento del día. Pero la presencia de una interacción significativa de segundo orden obliga a corregir esta conclusión.

Por un lado, el hecho de que la interacción de primer orden entre *tratamiento* y *momento* no sea significativa indica que no es posible afirmar que el efecto del tratamiento (la diferencia entre los grupos experimental y control) sea distinto en los tres momentos del día. Aunque esto podría llevar a pensar que el efecto del tratamiento es independiente del momento del día, la interacción de segundo orden indica que la interacción de primer orden entre *tratamiento* y *momento* no es la misma en las dos condiciones de hacinamiento. Este resultado ya se ha discutido más arriba (a propósito del efecto principal del tratamiento): aunque en la condición de alto hacinamiento el efecto del tratamiento es significativo en los tres momentos del día ( $p < 0,0005$  en los tres momentos), en la condición de bajo hacinamiento no puede afirmarse que exista efecto del tratamiento en ninguno de los tres momentos del día (mañana:  $p = 0,570$ ; tarde:  $p = 0,776$ ; noche:  $p = 0,069$ ).

Por otro lado, que la interacción de primer orden entre *hacinamiento* y *momento* no sea significativa indica que no es posible afirmar que la diferencia entre las condiciones de bajo y alto hacinamiento difiera entre los tres momentos del día. Aunque esto podría llevar a pensar que el efecto del hacinamiento es independiente del momento del día, la interacción de segundo orden está indicando que la interacción de primer orden

entre *hacinamiento* y *momento* no es la misma en los grupos experimental y control. Este resultado se ha discutido ya más arriba (a propósito del efecto principal del hacinamiento): mientras que en el grupo experimental el número de episodios agresivos es mayor en la condición de alto hacinamiento que en la de bajo hacinamiento ( $p < 0,0005$  en los tres momentos), en el grupo control no se observan diferencias entre las dos condiciones de hacinamiento en ninguno de los tres momentos (mañana:  $p = 0,478$ ; tarde:  $p = 0,887$ ; noche:  $p = 0,204$ ).

Parece, por tanto, que para interpretar correctamente los efectos presentes en un diseño de tres factores, no puede pasarse por alto una interacción de segundo orden significativa: su presencia matiza (incluso puede alterar) el significado del resto de efectos. Y para interpretar correctamente una interacción de segundo orden es necesario llevar a cabo los contrastes que permitan captar su naturaleza, si bien la información facilitada por otros contrastes (ya sean los *efectos simples de interacción*, los *efectos simples*, los *efectos simples simples* o, incluso, los *efectos principales*) puede resultar útil para entender el fenómeno observado en todos sus matices.

### 5.2.5. La sentencia **LMATRIX**

La sentencia *LMATRIX* no sólo sirve para comparar entre sí efectos simples de un diseño de dos factores (ver capítulo anterior). También permite llevar a cabo comparaciones entre los efectos simples de interacción en un diseño de tres factores. Para ello, siguiendo la misma lógica ya expuesta a propósito de la interacción de primer orden, es necesario añadir, a la sintaxis que se obtiene del cuadro de dialogo *MODELO LINEAL GENERAL > UNIVARIANTE*, la siguiente sentencia *LMATRRIX*:

/LMATRIX =

```
tto*hacinam*momento 1 -1 0 -1 1 0 -1 1 0 1 -1 0 ;
tto*hacinam*momento 1 0 -1 -1 0 1 -1 0 1 1 0 -1 ;
tto*hacinam*momento 0 1 -1 0 -1 1 0 -1 1 0 1 -1
```

La Tabla 5.32 muestra los resultados que se obtienen con esta estrategia. Puede comprobarse en ella que los valores de los contrastes y de los niveles críticos son idénticos a los obtenidos con la estrategia propuesta en el apartado anterior (ver Tabla 5.28).

**Tabla 5.32. Resultados de las comparaciones  $\psi_1$ ,  $\psi_2$  y  $\psi_3$  con LMATRIX**

Contraste	Valor del contraste	Error típico	Sig. (bilateral)
$L_1 = \psi_1$	-0,60	2,798	0,831
$L_2 = \psi_2$	-8,00	2,798	<b>0,006</b>
$L_3 = \psi_3$	-7,400	2,798	<b>0,011</b>

De nuevo, al comparar los efectos simples de interacción con esta estrategia, además de una valoración individual de cada comparación, se obtiene una valoración del efecto global de todas las comparaciones tomadas juntas. La Tabla 5.33 muestra esta valoración. Y ocurre que, al igual que en el caso de la interacción de primer orden estudiada en el apartado anterior, todos los resultados asociados a este efecto global (suma de cuadrados, grados de libertad, ..., nivel crítico) son idénticos a los del efecto global de la interacción (ver Tabla 5.27).

**Tabla 5.33. Efecto global de las comparaciones  $\psi_1$ ,  $\psi_2$  y  $\psi_3$**

Fuente	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Contraste	49,633	2	24,817	5,073	0,010
Error	234,800	48	4,892		



De nuevo, este resultado indica que las comparaciones que estamos proponiendo para estudiar el efecto de la interacción, no sólo se están refiriendo justamente a ese efecto, sino que lo agotan.

## 6. Conclusiones

1. La interacción entre factores en el análisis de varianza tiene un significado teórico inequívoco. Cuando se trata de *definir* la interacción (no de *analizarla* o *interpretarla*), la literatura estadística coincide en recoger ese significado tanto en el caso de dos como en el de tres factores (casos de los que nos hemos ocupado aquí).
2. A pesar del significado teórico inequívoco del concepto de interacción entre factores, en la investigación aplicada se utilizan con altísima frecuencia estrategias analíticas incorrectas que conducen a conclusiones erróneas: en nuestra revisión, se han encontrado interpretaciones correctas en menos del 10 % de las interacciones de primer orden y en menos del 5% de las de segundo orden.
3. La presencia de una tasa tan elevada de interpretaciones incorrectas de la interacción entre factores podría deberse a dos razones. En primer lugar, es común encontrar que los manuales de estadística y de análisis de datos no ofrezcan una explicación apropiada acerca de cómo proceder para analizar una interacción significativa. En segundo lugar, con los programas informáticos de análisis estadístico disponibles (el SPSS a la cabeza) no es posible obtener directamente (a través de cuadros de diálogo) las comparaciones que permiten analizar correctamente una interacción significativa.

4. Entre las estrategias analíticas incorrectas más utilizadas para interpretar una interacción significativa, además de las relacionadas con la omisión total o parcial del análisis del efecto de la interacción, existe una muy generalizada basada en el análisis de la significación individual de los *efectos simples*. Tanto en el caso de las interacciones de primer orden como en las de segundo orden, ha quedado demostrado con ejemplos que esta estrategia puede conducir a conclusiones erróneas porque las diferencias en la significación individual de los efectos simples no implica, necesariamente, diferencias entre efectos simples; y al revés, las diferencias entre efectos simples no implican diferencias en sus significaciones individuales.
5. La interpretación correcta de una interacción significativa exige plantear comparaciones lineales de un grado de libertad. Sólo de esta manera es posible aislar y agotar el significado de la interacción.
6. En los diseños factoriales completamente al azar es posible utilizar las prestaciones del programa estadístico SPSS para realizar las comparaciones necesarias para interpretar correctamente una interacción significativa. Para ello, se han propuesto dos estrategias alternativas pero equivalentes: los contrastes personalizados del procedimiento ONEWAY después de combinar los factores en una variable única; y la sentencia LMATRIX disponible con el procedimiento UNIANOVA.
7. Es necesario seguir investigando el problema para ofrecer a los investigadores en Psicología y ciencias afines soluciones fáciles de utilizar. También es necesario estudiar la interacción entre factores en los diseños de medidas repetidas para poder encontrar soluciones similares a las propuestas aquí para los diseños completamente al azar.

## 7. Referencias bibliográficas

- Abelson, R. P. y Prentice, D. A. (1997). Contrast tests of interaction hypotheses. *Psychological Bulletin*, 2, 315-328.
- Alarcón, R., y Blanca, M. J. (2000). Asimetría hemisférica en la dicotomía holística-analítica en tareas de atención focalizada. *Psicothema*, 12(Suppl2), 15-17.
- Algarabel, S., y Rosa, E. (2001). Medida de la integración del conocimiento estadístico y psicológico. *Psicothema*, 13(2), 302-309.
- Algarabel, S., Luciano, J. V., y Martínez, J. L. (2006). Inhibitory voluntary control of memory: Effect of stimulus onset asynchrony on reaction time to suppressed memories. *Psicológica*, 27(1), 57-77.
- Alonso, J., y Román, J. M. (2005). Prácticas educativas familiares y autoestima. *Psicothema*, 17(1), 76-82.
- Alvarado, J. M., y Santisteban, C. (2004). Compatibilidad de respuesta y el efecto simon. *Psicothema*, 16(2), 276-281.
- Alvarez, R. (1994) *Estadística multivariante y no paramétrica con SPSS aplicación a las ciencias de la salud*. Madrid: Consejería de Salud de la Comunidad de Madrid.
- Alvarez, M. (1998). *Introducción a la práctica de la estadística problemas reales resueltos con SPSS*. Bilbao: Universidad de Deusto.
- Amigo, I., Fernández, Á., González, A., y Herrera, J. (2002). Relajación muscular y monitorización ambulatoria de la presión arterial en la hipertensión esencial ligera. *Psicothema*, 14(1), 47-52.
- Ato, M. y Vallejo, G. (2007). *Diseños experimentales en Psicología*. Madrid: Pirámide.
- Betancor, V., Leyens, J., Rodríguez, A., y Quiles, M. N. (2003). Atribución diferencial al endogrupo y al exogrupo de las dimensiones de moralidad y eficacia: Un indicador de favoritismo endogrupal. *Psicothema*, 15(3), 407-413.
- Balluerka, N., Vergara, M. I., Arnau y Gras, J. (2002) *Diseños de investigación experimental en psicología modelos y análisis de datos mediante el SPSS 10.0*. Madrid, Prentice Hall.
- Bryman, A. (2001) *Quantitative data analysis with SPSS Release 10 for Windows: a guide for social scientists*. Hove (UK): Routledge; Philadelphia: Taylor y Francis.
- Camacho, J. (2004) *Estadística con SPSS (versión 12) para Windows*. Madrid: Rama.
- Campo, P., Morales, M., y Juan-Malpartida, M. (2000). Versiones españolas del test de recuerdo verbal selectivo. *Psicothema*, 12(Suppl2), 108-110.
- Cardenal, V., y Fierro, A. (2001). Sexo y edad en estilos de personalidad, bienestar personal y adaptación social. *Psicothema*, 13(1), 118-126.
- Chica, A. B., y Lupiáñez, J. (2004). Inhibición de retorno sin retorno de la atención. *Psicothema*, 16(2), 248-254.
- Chico, E. (2000). Búsqueda de sensaciones. *Psicothema*, 12(2), 229-235.

- Coakes, S. J. (2005) *SPSS Analysis without Anguish Version 12.0 for Windows*. John Wiley and Sons, Australia.
- Conejo, N. M., González, H., Cimadevilla, J. M., Vallejo, G., y Arias, J. L. (2003). Maduración de los astrocitos del hipocampo de la rata: Posibles implicaciones conductuales. *Psicothema*, 15(2), 216-220.
- Docampo, M. M. (2002). Diferencias de género en las explicaciones causales de adolescentes. *Psicothema*, 14(3), 572-576.
- Durán, E., Vargas, J. P., Salas, C., y Papini, M. R. (2000). Effect of telencephalic ablation on appetitive instrumental learning in goldfish (*carassius auratus*). *Psicothema*, 12(4), 520-524.
- Everitt, B. S. y Howell, D. C. (2005). *Encyclopedia of statistics in behavioral science* (vol 2, págs. 929-933). Chichester, Sussex: Wiley.
- Fernández, R., Secades, R., Terrados, N., García, E., y García, J. M. (2004). Efecto de la hipnosis y la terapia de aceptación y compromiso (ACT) en la mejora de la fuerza física en piragüistas. *International Journal of Clinical and Health Psychology*, 4(3), 481-493.
- Ferrán, M. (2001) *SPSS para Windows : análisis estadístico*. Madrid: Osborne McGraw-Hill.
- Field, A. (2005). *Discovering statistics using SPSS* (2ª ed.). Thousand Oaks, CA: Sage.
- Games, P. A. (1973). Type IV errors revised. *Psychological Bulletin*, 80, 304-307.
- García, J., y Arias-Gundín, O. (2004). Intervención en estrategias de revisión del mensaje escrito. *Psicothema*, 16(2), 194-202.
- Gerber, S. B. Voelkl K. (2005) *Using SPSS for windows Data analysis and graphics* (2nd ed.). New Cork: Springer.
- González, J. A., Núñez, J. C., González, S., Álvarez, L., Rocés, C., García, M., y col. (2000). Autoconcepto, proceso de atribución causal y metas académicas en niños con y sin dificultades de aprendizaje. *Psicothema*, 12(4), 548-556.
- González, R. M. (2003). Propuesta de intervención en los procesos cognitivos y estructuras textuales en niños con DAE. *Psicothema*, 15(3), 458-463.
- Gracia, F. J., Arcos, J. L., y Caballer, A. (2000). Influencia de la presión temporal en el trabajo en grupo en función del tipo de tarea y del canal de comunicación. *Psicothema*, 12(2), 241-246.
- Guzmán, R., Jiménez, J. E., Ortiz, M. R., Hernández, I., Estévez, A., Rodrigo, M., y col. (2004). Evaluación de la velocidad de nombrar en las dificultades de aprendizaje de la lectura *Psicothema*, 16(3), 442-447.
- Hays, W. L. (1994). *Statistics* (5ª ed.). Belmont, CA: Wadsworth.
- Ho, R. (2006) *Handbook of univariate and multivariate data analysis and interpretation with SPSS*. Boca Raton (US): Chapman y Hall/CRC.
- Howell, D. C. (2002). *Statistical methods for psychology* (5ª ed.). Boston: PWS-Kent Publishing Company.
- Jaccard, J. (1998). *Interaction effects in factorial analysis of variance*. Thousand Oaks, CA: Sage.
- Keppel, G. y Wickens, Th. D. (2004). *Design and analysis: A researcher's handbook* (4ª ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Kinney, P. R., y Gray C. D. (2006) *SPSS 14 made simple*. New Cork: Psychology Press.

- Kirk, R. E. (1995). *Experimental design. Procedures for the behavioral sciences* (3ª ed.). Belmont, CA: Brooks/Cole.
- Landau, S., y Everitt, B.S. (2004) *A handbook of statistical analyses using SPSS*. Boca Raton (US): Chapman y Hall/CRC.
- León, O. G. y Montero, I. (2001). Cómo explicar el concepto de interacción sin estadística: análisis gráfico de todos los casos posibles en un diseño 2x2. *Psicothema*, 13, 159-165.
- Levin, J. R. y Marascuilo, L. A. (1972). Type IV errors and interactions. *Psychological Bulletin*, 78, 368-374.
- Levin, J. R. y Marascuilo, L. A. (1973). Type IV errors and Games. *Psychological Bulletin*, 80, 308-309.
- Marascuilo, L. A. y Levin, J. R. (1970). Appropriate post hoc comparisons for interaction and nested hypotheses in analysis of variance designs: the elimination of type IV errors. *American Educational Research Journal*, 7, 397-421.
- Marcos, J. L., y Redondo, J. (2005). Facilitation and interference of the automatic information processing on a reaction time task to threat-relevant stimuli. *Psicothema*, 17(2), 332-337.
- Martín, P. J. (2006) *Prácticas de tratamiento estadístico de datos con el programa SPSS para Windows Aplicaciones en el área de Ciencia y Tecnología de alimentos*. Madrid, Consejo Superior de Investigaciones Científicas.
- Martínez, J., Muñiz, J., y García, E. (2000). Mejora de las puntuaciones de los tests mediante el entrenamiento. *Psicothema*, 12(Supl2), 363-367.
- Martínez, M.A. (2007) En Martínez, M.A., Sánchez, A., Faulin, F.J. (Ed.) *Estadística amigable*. Madrid: Díaz de Santos.
- Maxwell, S. E. y Delaney, H. D. (2004). *Designing experiments and analyzing data. A model comparison perspective* (2ª ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Meyer, D. L. (1991). Misinterpretation of interactions effects: A reply to Rosnow and Rosenthal. *Psychological Bulletin*, 110, 571-573.
- Miller, R. L. (2002) En Campling, J. (Ed.) *SPSS for social scientists*. Houndmills, Basingstoke, Hampshire New York: Palgrave MacMillan.
- Milton, J.S. (2007) *Estadística para biología y ciencias de la salud* (3ª ed. ampliada). Madrid, Interamericana- McGraw-Hill.
- Miñarro, J., Rodríguez, M., Pérez, N., Restall, C. J., y Brain, P. F. (2000). Developmental and neurochemical effects of early postnatal exposure to flumazenil in female and male mice. *Psicothema*, 12(4), 648-653.
- Monsalve, A., y Cueto, F. (2001). Asimetría hemisférica en el reconocimiento de palabras: Efectos de frecuencia e imaginabilidad. *Psicothema*, 13(1), 24-28.
- Nuevo, R., y Montorio, I. (2005). Eficacia de un tratamiento cognitivo-conductual para la preocupación excesiva en personas mayores. *International Journal of Clinical and Health Psychology*, 5(2), 215-226.
- Ottensbacher, K. J. (1991). Interpretation of Interaction in Factorial Analysis of Variance Design. *Statistics in Medicine*, 10, 1565-1571.
- Pacheco, V., Flores, C., González, F., Canales, C., y Carpio, C. (2005). Efectos de la consistencia e inconsistencia de las relaciones intrusivo-reforzador y muestra-reforzador en igualación a la muestra. *Psicothema*, 17(1), 118-122.
- Pardo, A. y San Martín, R. (1998). *Análisis de datos en psicología II* (2ª ed.). Madrid: Pirámide.

- Pardo, A. y Ruiz, M. A. (2005). *Análisis de datos con SPSS 13 Base*. Madrid: McGraw-Hill.
- Pascual, J. (1998). Diseño entre grupos. En M. T. Anguera, J. Arnau, M. Ato, R. Martínez, J. Pascual y G. Vallejo (Ed.), *Métodos de investigación en psicología*, Madrid: Síntesis (págs. 73-112).
- Pedhazur, E. J. y Pedhazur, L. (1991). *Measurement, design and analysis. An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pérez, A. M., Sanjuán, P., y Bermúdez, J. (2002). Control personal y situacional y reactividad cardiovascular en tareas de estrés físico. *Psicothema*, 14(3), 583-590.
- Pérez, C. (2004) *Técnicas de análisis multivariante de datos aplicaciones con SPSS*. Madrid, Pearson: Prentice Hall.
- Pérez, C. (2005a) *Métodos estadísticos avanzados con SPSS*. Madrid, Thomson Editores.
- Pérez, C. (2005b) *Técnicas estadísticas con SPSS® 12 aplicaciones al análisis de datos*. Madrid: Pearson Educación.
- Petty, R. E., Fabrigar, L. R., Wegener, D. T. y Priester, J. R. (1996). Understanding data when interactions are present or hypothesized. *Psychological Science*, 7, 247-252.
- Pitarque, A., Algarabel, S., Dasí, C., y Ruiz, J. C. (2003). Olvido dirigido de falsas memorias: Podemos olvidar intencionadamente una falsa memoria?. *Psicothema*, 15(1), 6-11.
- Prieto, M. D., López, O., Bermejo, M. R. B., Renzulli, J., y Castejón, J. L. (2002). Evaluación de un programa de desarrollo de la creatividad. *Psicothema*, 14(2), 410-414.
- Recarte, M. Á, Conchillo, Á, y Nunes, L. M. (2005). Estimation of arrival time in vehicle and video. *Psicothema*, 17(1), 112-117.
- Rodríguez, M. J. y Mora, R. (2001) *Estadística informática. Casos y ejemplos con el SPSS*. Alicante: Universidad.
- Rodríguez, R., y Moya, M. (2003). La diferencia percibida en valores como proceso vinculado a las relaciones intergrupales de payos y gitanos. *Psicothema*, 15(2), 176-182.
- Rodríguez, G., y Alonso, G. (2004). Latent inhibition as a function of US intensity in a two-stage CER procedure. *Psicológica*, 25(2), 207-216.
- Rodríguez, A., Betancor, V., Rodríguez, R., Quiles, M. N., Delgado, N., y Coello, E. (2005). El efecto de las identidades nacionales con distintos niveles de inclusividad en el prejuicio hacia exogrupos. *Psicothema*, 17(3), 441-446.
- Rosnow, R. L. y Rosenthal, R. (1989a). Definition and interpretation of interaction effects. *Psychological Bulletin*, 105, 143-146.
- Rosnow, R. L. y Rosenthal, R. (1989b). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist*, 44, 1276-1284.
- Rosnow, R. L. y Rosenthal, R. (1991). If you're looking at the cell means, you're not looking only at the interaction (unless all main effects are zero). *Psychological Bulletin*, 110, 574-576.
- Rosnow, R. L. y Rosenthal, R. (1995). "Some things you learn aren't so": Cohen's paradox, Asch's paradigm, and the interpretation of interaction. *Psychological Science*, 6, 3-9.
- Rosnow, R. L. y Rosenthal, R. (1996). Contrast and interactions redux: Five easy pieces. *Psychological Science*, 7, 253-257.
- Ruiz, J. A., y Sánchez, C. (2001). Depresión y memoria: Es la información congruente con el estado de ánimo más accesible?. *Psicothema*, 13(2), 193-196.

- Sánchez, M., Ruiz, E., Pérez, N., Fernández, M. C., Cobos, P., y Vila, J. (2002). Modulación emocional de los reflejos defensivos mediante visualización de imágenes afectivas. *Psicothema*, 14(4), 702-707.
- Sánchez, J. P., Martínez, J. M., Román, F., y Torrente, G. (2006). The effect of content and physical properties of affective pictures on emotional responses. *The Spanish Journal of Psychology*, 9(2), 145-153.
- Santostefano, S., Quiroga, M. A., y Rooney, S. (2001). Life stressors and cognitive styles in children. *The Spanish Journal of Psychology*, 4(1), 37-47.
- Sierra, B., Falces, C., y Briñol, P. (2002). Recuerdo sobre situaciones reales basadas en guiones: Relevancia y tipicidad de las acciones. *Psicothema*, 14(4), 776-782.
- Soriano, M. F., Macizo, P., y Bajo, T. (2004). Diferencias individuadas en tareas de interferencia episódica y semántica. *Psicothema*, 16(2), 187-193.
- Tornay, F. J., y Milán, E. G. (2001). Diferente participación de los mecanismos de control en el cambio entre tareas regular frente al cambio al azar. *Psicothema*, 13(1), 111-117.
- Vera, P. E., Pérez, V., Moreno, E., y Allende, F. (2004). Diferencias en variables psicosociales en sujetos VIH homosexuales y heterosexuales. *International Journal of Clinical and Health Psychology*, 4(1), 55-67.
- Verdugo, M. Á., y Sabeh, E. N. (2002). Evaluación de la percepción de calidad de vida en la infancia. *Psicothema*, 14(1), 86-91.
- Vicente, M. L. (2005) *Diseño de experimentos. Soluciones con SAS y SPSS*. Madrid: Pearson-Prentice Hall.
- Visauta, B., y Martorí i Cañas, J.C. (2003) *Análisis estadístico con SPSS para Windows: II Estadística multivariante* (2ª ed.) Madrid: McGraw-Hill.
- Winer, B. J., Brown, D. R. y Michels, K. M. (1991). *Statistical principles in experimental design* (3ª ed.). New York: McGraw-Hill.
- Zuckerman, M., Hodgins, H. S., Zuckerman, M., y Rosenthal, R. (1993). Contemporary Issues in the analysis of Data: A Survey of 551 Psychologists. *Psychological Science*, 4, 49-53.









## 8. ANEXOS

### ANEXO 1. Artículos revisados

- Aciego de Mendoza, R., Domínguez, R., y Hernández, P. (2003). Evaluación de la mejora en valores de realización personal y social en adolescentes que han participado en un programa de intervención. *Psicothema*, 15(4), 589-594.
- Ackel, A., y Lippi, S. (2006). Influence of the instructions on the performance and establishment of memorization strategies in space judgments. *The Spanish Journal of Psychology*, 9(2), 312-320.
- Acosta, M. V., y Gómez, G. (2003). Insatisfacción corporal y seguimiento de dieta. una comparación transcultural entre adolescentes de España y México. *International Journal of Clinical and Health Psychology*, 3(1), 9-21.
- Aguado, L., Pierna, M., y Saugar, C. (2005). Affective priming with associatively acquired valence. *Psicológica*, 26(2), 261-279.
- Aivar, M. P., Brenner, E., y Smeets, J. B. J. (2005). Correcting slightly less simple movements. *Psicológica*, 26(1), 61-79.
- Alarcón, R., y Blanca, M. J. (2000). Asimetría hemisférica en la dicotomía holística-analítica en tareas de atención focalizada. *Psicothema*, 12(Suppl2), 15-17.
- Alcalá, V., Camacho, M., Giner, D., Giner, J., y Ibáñez, E. (2006). Afectos y género. *Psicothema*, 18(1), 143-148.
- Alegre, J. (2000). El procesamiento visual humano, efectos de la información previa sobre la detección de estímulos poco contrastados. *Psicothema*, 12(Suppl2), 22-24.
- Algarabel, S., y Rosa, E. (2001). Medida de la integración del conocimiento estadístico y psicológico. *Psicothema*, 13(2), 302-309.
- Algarabel, S., Pitarque, A., y Dasí, C. (2002). Efecto del tipo de información sobre el esfuerzo cognitivo y el estado de consciencia asociado al recuerdo. *Psicothema*, 14(2), 393-398.
- Algarabel, S., Luciano, J. V., y Martínez, J. L. (2006). Inhibitory voluntary control of memory: Effect of stimulus onset asynchrony on reaction time to suppressed memories. *Psicológica*, 27(1), 57-77.
- Alonso, J., y Román, J. M. (2005). Prácticas educativas familiares y autoestima. *Psicothema*, 17(1), 76-82.
- Alvarado, J. M., y Santisteban, C. (2004). Compatibilidad de respuesta y el efecto simon. *Psicothema*, 16(2), 276-281.
- Álvarez, A., Blanco, M., y Leirós, L. (2002). Influencia de la simetría y la curvilinealidad en el procesamiento de estímulos cerrados. *Psicothema*, 14(3), 597-604.

- Álvarez, M., y Villamarín, F. (2004). El papel de la autoeficacia en el entrenamiento para controlar la frecuencia cardíaca durante pruebas de esfuerzo. *Psicothema*, 16(1), 50-57.
- Álvarez, J. L. (2005). Las metas en la reducción del prejuicio automático. *Psicothema*, 17(1), 71-75.
- Amigo, I., Fernández, Á., González, A., y Herrera, J. (2002). Relajación muscular y monitorización ambulatoria de la presión arterial en la hipertensión esencial ligera. *Psicothema*, 14(1), 47-52.
- Arend, I., y Botella, J. (2002). Emotional stimuli reduce the attentional blink in sub-clinical anxious subjects. *Psicothema*, 14(2), 209-214.
- Ato, E., González, C., Carranza, J. A., y Ato, M. (2004). Malestar y conductas de autorregulación ante la situación extraña en niños de 12 meses de edad. *Psicothema*, 16(1), 1-6.
- Augusto, J. M., Adrián, J. A., Alegría, J., y Martínez de Antóñana, R. (2002). Dificultades lectoras en niños con sordera. *Psicothema*, 14(4), 746-753.
- Avero, P., y Calvo, M. G. (2006). Affective priming with pictures of emotional scenes: The role of perceptual similarity and category relatedness. *The Spanish Journal of Psychology*, 9(1), 10-18.
- Aznar, J. A., Amador, J. A., Freixa, M., y Turbany Oset, J. (2000). Consumo atencional en la estimación de la profundidad retrovisual. *Psicothema*, 12(1), 71-78.
- Baile, J. I., Guillén, F., y Garrido, E. (2002). Insatisfacción corporal en adolescentes medida con el body shape questionnaire (BSQ): Efecto del anonimato, el sexo y la edad. *International Journal of Clinical and Health Psychology*, 2(3), 439-450.
- Bajo, M. T., Puerta-Melguizo, M. C., y Macizo, P. (2003). The locus of semantic interference in picture naming. *Psicológica*, 24(1), 31-55.
- Ballesteros, S., Reales, J. M., y Manga, D. (2000). Effects of type of design (blocked vs. randomized) on stroop and emotional stroop tasks. *Psicothema*, 12(Suppl2), 60-63.
- Ballesteros, M. A., Gallo, M., y Maldonado, A. (2001). Detección de aversiones gustativas inducidas por estímulos incondicionados débiles como la rotación. *Psicológica*, 22(2), 217-234.
- Ballesteros, S., Reales, J. M., García, E., y Manga, D. (2006). Selective attention affects implicit and explicit memory for familiar pictures at different delay conditions. *Psicothema*, 18(1), 88-99.
- Baños, R. M., Quero, S., y Botella, C. (2005). Sesgos atencionales en la fobia social medidos mediante dos formatos de la tarea stroop emocional (de tarjetas y computerizado) y papel mediador de distintas variables clínicas. *International Journal of Clinical and Health Psychology*, 5(1), 23-42.
- Baqués, J., Sáiz, D., y Sáiz, M. (2003). Medida de la memoria retrógrada a través de un cuestionario sobre recuerdo de personajes y sucesos del pasado. *Psicothema*, 15(3), 401-406.
- Barberá, E., y Cala, M. J. (2004). Creencias estereotipadas sobre ciclicidades. *Psicothema*, 16(1), 14-21.
- Bermejo, V. (2005). Microgénesis y cambio cognitivo: Adquisición del cardinal numérico. *Psicothema*, 17(4), 559-562.
- Bernardo, I., Bernardo, A., y Herrero, J. (2005). Nuevas tecnologías y educación especial. *Psicothema*, 17(1), 64-70.

- Bernedo, I. M., Fuentes, M. J., y Fernández, M. (2005). Percepción del grado de conflicto en familias adoptivas y no adoptivas. *Psicothema*, 17(3), 370-374.
- Besteiro, J. L., y García, E. (2000). Utilización de la hipnosis en el tratamiento de la depresión mayor. *Psicothema*, 12(4), 557-560.
- Betancor, V., Rodríguez, A., Rodríguez, A., y Sánchez, J. (2001). La activación del estereotipo bajo anestesia general reduce la ambigüedad de la información social. *Psicothema*, 13(2), 318-323.
- Betancor, V., Leyens, J., Rodríguez, A., y Quiles, M. N. (2003). Atribución diferencial al endogrupo y al exogrupo de las dimensiones de moralidad y eficacia: Un indicador de favoritismo endogrupal. *Psicothema*, 15(3), 407-413.
- Betancor, V., Rodríguez, A., Quiles, M. N., y Rodríguez, R. (2005a). Relación de la infrahumanización del exogrupo con los procesos de inferencia y memoria. *Psicothema*, 17(3), 447-452.
- Betancor, V., Rodríguez, A., Rodríguez, R., Leyens, J., y Quiles, M. N. (2005b). El efecto del estatus en la atribución de las dimensiones estereotípicas de sociabilidad y competencia. *Psicothema*, 17(2), 297-302.
- Blanca, M. J., Luna, R., López, D., Zalabardo, C., y Rando, B. (2001). Características de los estímulos y de la tarea en el procesamiento de los rasgos global y local. *Psicológica*, 22(2), 267-291.
- Botella, J., Barriopedro, M. I., y Joula, J. F. (2002). Temporal interactions between target and distractor processing: Positive and negative priming effects. *Psicológica*, 23(2), 371-400.
- Botella, J., Arend, I., y Suero, M. (2004). Illusory conjunctions in the time domain and the resulting time-course of the attentional blink. *The Spanish Journal of Psychology*, 7(1), 63-68.
- Brandimonte, M. A., Ferrante, D., Feresin, C., y Delbello, R. (2001). Dissociating prospective memory from vigilance processes. *Psicológica*, 22(1), 97-113.
- Bretón, J., y Buela, G. (2005). Evaluación del efecto de las campañas publicitarias de prevención de VIH/SIDA en adolescentes. *Psicothema*, 17(4), 590-596.
- Briñol, P., Sierra, B., Falces, C., Becerra, A., y Froufe, M. (2000). La eficacia relativa del efecto de mera exposición y del condicionamiento clásico en la formación de preferencias. *Psicothema*, 12(4), 586-593.
- Briñol, P., Horcajo, J., Becerra, A., Falces, C., y Sierra, B. (2002). Cambio de actitudes implícitas. *Psicothema*, 14(4), 771-775.
- Briñol, P., Becerra, A., Gallardo, I., Horcajo, J., y Valle, C. (2004a). Validación del pensamiento y persuasión. *Psicothema*, 16(4), 606-610.
- Briñol, P., Gallardo, I., Horcajo, J., de la Corte, L., Valle, C., y Díaz, D. (2004b). Afirmación, confianza y persuasión. *Psicothema*, 16(1), 27-31.
- Briones, E., y Tabernero, C. (2005). Formación cooperativa en grupos heterogéneos. *Psicothema*, 17(3), 396-403.
- Bueno, M., y Álvarez, R. (2001). El efecto de las duraciones del intervalo entre ensayos y entre estímulos en el condicionamiento pavloviano apetitivo en ratas. *Psicológica*, 22(2), 205-215.
- Bykov, A. T., Malyarenko, T. N., Malyarenko, Y. E., Terentjev, V. P., y Dyuzhikov, A. A. (2006). Conscious and unconscious sensory inflows allow effective control of the functions of the human brain and heart at the initial ageing stage. *The Spanish Journal of Psychology*, 9(2), 201-218.

- Cáceres, A., y Cáceres, J. (2006). Violencia en relaciones íntimas en dos etapas evolutivas. *International Journal of Clinical and Health Psychology*, 6(2), 271-284.
- Callejas, A., Correa, Á., Lupiáñez, J., y Tudela, P. (2003). Normas asociativas intracategoriales para 612 palabras de seis categorías semánticas en español. *Psicológica*, 24(2), 185-200.
- Calvo, M. G., y Castillo, M. D. (2005). Processing of threat-related information outside the focus of visual attention. *The Spanish Journal of Psychology*, 8(1), 3-11.
- Camarero, F., Martín del Buey, F., y Herrero, J. (2000). Estilos y estrategias de aprendizaje en estudiantes universitarios. *Psicothema*, 12(4), 615-622.
- Campo, P., Morales, M., y Juan-Malpartida, M. (2000). Versiones españolas del test de recuerdo verbal selectivo. *Psicothema*, 12(Suppl2), 108-110.
- Cándido, A., Perales, J. C., Catena, A., Maldonado, A., Guadarrama, L., Beltrán, R., y col. (2006). Efectos de la inducción emocional en el aprendizaje causal. *Psicológica*, 27(2), 243-267.
- Cano, F. (2000). Diferencias de género en estrategias y estilos de aprendizaje. *Psicothema*, 12(3), 360-367.
- Carbonero, M. Á., y Merino, E. (2004). Autoeficacia y madurez vocacional. *Psicothema*, 16(2), 229-234.
- Cardenal, V., y Fierro, A. (2001). Sexo y edad en estilos de personalidad, bienestar personal y adaptación social. *Psicothema*, 13(1), 118-126.
- Catena, A., Perales, J. C., y Maldonado, A. (2004). Judgment frequency effects in generative and preventative causal learning. *Psicológica*, 25(1), 67-85.
- Cecchini, J. A., Montero, J., y Peña, J. V. (2003). Repercusiones del programa de intervención para desarrollar la responsabilidad personal y social de hellison sobre los comportamientos de fair-play y el auto-control. *Psicothema*, 15(4), 631-637.
- Cecchini, J. A., González, C., Carmona, A. M., y Contreras, O. (2004). Relaciones entre clima motivacional, orientación de meta, motivación intrínseca, auto-confianza, ansiedad y estado de ánimo en jóvenes deportistas. *Psicothema*, 16(1), 104-109.
- Chica, A. B., y Lupiáñez, J. (2004). Inhibición de retorno sin retorno de la atención. *Psicothema*, 16(2), 248-254.
- Chico, E. (2000). Búsqueda de sensaciones. *Psicothema*, 12(2), 229-235.
- Cimadevilla, J. M., Conejo, N. M., Miranda, R., y Arias, J. L. (2004). Sex differences in the morris water maze in young rats: Temporal dimensions. *Psicothema*, 16(4), 611-614.
- Cobos, P., García, C., Ríus, F., y Vila, J. (2002). Modulación emocional de la respuesta de sobresalto. *Psicothema*, 14(1), 106-111.
- Conchillo, Á., Hernández, M. J., Recarte, M. A., y Ruiz, T. (2000). La psicofísica de la velocidad en el contexto de la conducción real de automóviles. *Psicothema*, 12(Suppl2), 152-156.
- Conchillo, Á., Recarte, M. Á., Nunes, L., y Ruiz, T. (2006). Comparing speed estimations from a moving vehicle in different traffic scenarios: Absence versus presence of traffic flow. *The Spanish Journal of Psychology*, 9(1), 32-37.
- Conde, M., Menéndez, F. J., y López, A. (2000) Estabilización de la actividad electrodermal y ciertos parámetros respiratorios. *Psicothema*, 12(3), 426-434.

- Conejo, N. M., González, H., Cimadevilla, J. M., Vallejo, G., y Arias, J. L. (2003). Maduración de los astrocitos del hipocampo de la rata: Posibles implicaciones conductuales. *Psicothema*, 15(2), 216-220.
- Conejo, N. M., López, M., González, H., Cantora, R., Begega, A., López, L., y col. (2005). Brain metabolism after extended training in a fear conditioning task. *Psicothema*, 17(4), 563-568.
- Cuadrado, I. (2004). Valores y rasgos estereotípicos de género de mujeres líderes. *Psicothema*, 16(2), 270-275.
- Cuevas, L. M., Nuñez, J. C., Rodríguez, F. J., y González, N. (2002). Elaboración y evaluación de un programa de mejora de la comprensión oral. *Psicothema*, 14(2), 293-299.
- Da Silva, J. A., Matsushima, E. H., Aznar-Casanova, J. A., y Ribeiro-Filho, N. P. (2006). Distance perception in a natural outdoor setting: Is there a developmental trend to overconstancy? *The Spanish Journal of Psychology*, 9(2), 285-294.
- Damián, M. (2000). Escala del desarrollo comunicativo-social aplicada a niños españoles y mexicanos. *Psicothema*, 12(Suppl2), 163-165.
- de Brugada, I., y Aguado, L. (2000). El efecto de preexposición al EI en la aversión gustativa condicionada: Papel relativo del contexto y de las claves de inyección. *Psicológica*, 21(1-2), 23-37.
- de Gracia, M., y Marcó, M. (2000). Efectos psicológicos de la actividad física en personas mayores. *Psicothema*, 12(2), 285-292.
- de Paúl, J., Pérez, A., Paz, P. M., Alday, N., y Moco-ro, I. (2002). Recuerdos de maltrato infantil en maltratadores y potencial de maltrato en víctimas de maltrato físico y abuso sexual. *Psicothema*, 14(1), 53-62.
- del Barrio, J. A., y Gutiérrez, J. N. (2000). Diferencias en el estilo de aprendizaje. *Psicothema*, 12(2), 180-186.
- del Campo, A., y López, F. (2006). Evaluación de un programa de prevención de abusos sexuales a menores en educación primaria. *Psicothema*, 18(1), 1-8.
- del Pino, A., Gaos, M. T., Dorta, R., y García, M. (2004). Eficacia de un programa cognitivo-conductual para modificar conductas prono-coronarias. *Psicothema*, 16(4), 654-660.
- del Pino, A., Gaos, M. T., Dorta, R., y García, M. (2005). Modification of coronary-prone behaviors in coronary patients of low socio-economic status. *The Spanish Journal of Psychology*, 8(1), 68-78.
- del Río, D., y López, R. (2006). Efectos de la memoria operativa y de una carga de procesamiento en la comprensión de oraciones. *Psicológica*, 27(1), 79-95.
- Díaz, A., y Pérez, J. (2003). Atención y nivel de actividad hacia personas y objetos físicos durante el primer año de vida: el papel del temperamento. *Psicothema*, 15(2), 234-240.
- Díaz, E., de la Casa, L. G., Ruiz, G., y Baeyens, F. (2004). Aprendizaje sabor-sabor en la adquisición de preferencias gustativas. *Psicológica*, 25(2), 135-146.
- Docampo, M. M. (2002). Diferencias de género en las explicaciones causales de adolescentes. *Psicothema*, 14(3), 572-576.
- Durán, E., Vargas, J. P., Salas, C., y Papini, M. R. (2000). Effect of telencephalic ablation on appetitive instrumental learning in goldfish (*carassius auratus*). *Psicothema*, 12(4), 520-524.
- Elosua, P., López, A., y Egaña, J. (2000). Fuentes potenciales de sesgo en una prueba de aptitud numérica. *Psicothema*, 12(3), 376-382.



- Elosua, P., López, A., y Egaña, J. (2000). Fuentes potenciales de sesgo en una prueba de aptitud numérica. *Psicothema* 12(3), 376-382.
- Estevez, A., y Calvo, M. G. (2002). Context constraints, prior vocabulary knowledge and on-line inferences in reading. *Psicothema*, 14(2), 357-362.
- Estévez, A. F., y Fuentes, L. J. (2003). Differential outcomes effect in four-year-old children. *Psicológica*, 24(2), 159-167.
- Estrada, C., Yzerbyt, V., y Seron, E. (2004). Efecto del esencialismo psicológico sobre las teorías ingenuas de las diferencias grupales. *Psicothema*, 16(2), 181-186.
- Ettxebarria, I., Isasi, X., y Pérez, J. (2002). Naturaleza interpersonal de los eventos que provocan culpa. diferencias de edad y género. *Psicothema*, 14(4), 783-787.
- Falces, C., Briñol, P., Sierra, B., Becerra, A., y Alíer, E. (2001). Validación de la escala de necesidad de cognición y su aplicación al estudio del cambio de actitudes. *Psicothema*, 13(4), 622-628.
- Fernández, Á, y Alonso, M. A. (2001). The relative value of environmental context reinstatement in free recall. *Psicológica*, 22(2), 253-266.
- Fernández, J., y Merino, H. (2002). Sesgos de memoria implícita para información emocional en depresión subclínica *Psicothema*, 14(4), 795-801.
- Fernández, R., Secades, R., Terrados, N., García, E., y García, J. M. (2004a). Efecto de la hipnosis y la terapia de aceptación y compromiso (ACT) en la mejora de la fuerza física en piragüistas. *International Journal of Clinical and Health Psychology*, 4(3), 481-493.
- Fernández, J., Merino, H. M., y Pardo, J. L. (2004b). Sesgos de procesamiento de información emocional en individuos sociotrópicos. *Psicothema*, 16(2), 235-240.
- Fernández, J. L., Rivero, M. P., Quiros, P., Conde, M., y Guerra, B. (2005). Entre la percepción visual de lo posible y lo imposible en pacientes con esquizofrenia. *Psicothema*, 17(4), 607-613.
- Ferraces, M. J., Andrade, E. M., y Arce, C. (2000). Diferencias en la evaluación del servicio de autobús urbano desde la perspectiva del usuario. *Psicothema*, 12(suppl.2), 211-213.
- Ferré, P. (2002). Advantage for emotional words in immediate and delayed memory tasks: Could it be explained in terms of processing capacity? *The Spanish Journal of Psychology*, 5(2), 78-89.
- Figueiredo, B., Pacheco, A., Costa, R., y Magarinho, R. (2006). Gravidez na adolescência: Das circunstâncias de risco às circunstâncias que favorecem a adaptação à gravidez. *International Journal of Clinical and Health Psychology*, 6(1), 97-125.
- Fontecilla, P., y Calvete, E. (2003). Racionalidad y personas mayores. *Psicothema*, 15(3), 369-374.
- Fraga, I., García, J., y Acuña, J. C. (2005). La desambiguación de oraciones de relativo en gallego: Nueva evidencia de adjunción alta en lenguas romances. *Psicológica*, 26(2), 243-260.
- Frenz, H., y Lappe, M. (2006). Visual distance estimation in static compared to moving virtual scenes. *The Spanish Journal of Psychology*, 9(2), 321-331.
- Froufe, M., y Schwartz, C. (2001). Subliminal messages for increasing self-esteem: Placebo effect. *The Spanish Journal of Psychology*, 4(1), 19-25.
- Funes, M. J., y Lupiáñez, J. (2003). La teoría atencional de posner: Una tarea para medir las funciones atencionales de orientación, alerta y control cognitivo y la interacción entre ellas. *Psicothema*, 15(2), 260-266.

- Furnham, A., y Chamorro, T. (2005). Estimating one's own and one's relatives' multiple intelligence: A study from argentina. *The Spanish Journal of Psychology*, 8(1), 12-20.
- Galera, C., von Grünau, M., y Panagopoulos, A. (2005). Automatic focusing of attention on object size and shape. *Psicológica*, 26(1), 147-160.
- Garaigordobil, M. (2004). Intervención psicológica en la conducta agresiva y antisocial con niños. *Psicothema*, 16(3), 429-435.
- García, M. d. P., y Becoña, E. (2000a). Evaluation of the amount of therapist contact in a smoking cessation program. *The Spanish Journal of Psychology*, 3(1), 28-36.
- García, F., Colom, R., Lora, S., Rivas, M., y Traver, V. (2000b). Valoración de "filosofía para niños": Un programa de enseñar a pensar. *Psicothema*, 12(2), 207-211.
- García, P., Gómez, L., y Canto, J. M. (2001). Reacción de celos ante una infidelidad: Diferencias entre hombres y mujeres y características del rival. *Psicothema*, 13(4), 611-616.
- García, A., Vila, J., Reyes del Paso, G., y Pérez, M. N. (2002). Efectos de la frustración y/o el hostigamiento en la modulación emocional de la respuesta cardíaca de defensa. *Psicothema*, 14(2), 450-455.
- García, A., y Rosas, J. M. (2003a). Recuperación de la relación clave-consecuencia por el cambio de contexto después de la interferencia en aprendizaje causal. *Psicológica*, 24(2), 243-269.
- García, A., y Rosas, J. M. (2003b). The role of the number of cues on retroactive interference in human predictive learning. *Psicológica*, 24(2), 271-287.
- García, J., y Arias-Gundín, O. (2004a). Intervención en estrategias de revisión del mensaje escrito. *Psicothema*, 16(2), 194-202.
- García, L. M., Capilla, A., García, O., Luque, J., Senderek, K., Conejo, N. M., y col. (2004b). Alcohol tolerance in rats submitted to different periods of chronic and acute ethanol intake. *Psicothema*, 16(2), 211-216.
- García, L. J., Olivares, J., y Hidalgo, M. D. (2005a). A pilot study on sensitivity of outcome measures for treatments of generalized social phobia in spanish adolescents. *International Journal of Clinical and Health Psychology*, 5(2), 385-392.
- García, E., Fernández, P., y Rico, R. A. (2005b). Género y sexo como variables moduladoras del comportamiento sexual en jóvenes universitarios. *Psicothema*, 17(1), 49-56.
- García de la Banda, G., Martínez, Á, Riesco, M., y Pérez, G. (2004). La respuesta de cortisol ante un examen y su relación con otros acontecimientos estresantes y con algunas características de personalidad. *Psicothema*, 16(2), 294-298.
- Gawryszewski, L. G., Carreiro, L. R. R., y Magalhães, F. V. (2005). Early and late inhibitions elicited by a peripheral visual cue on manual response to a visual target: Are they based on cartesian coordinates? *Psicológica*, 26(1), 121-137.
- Gil-Burmann, C., Peláez, F., y Sánchez, S. (2002). Elección de pareja estable a través de anuncios de periódico. *Psicothema*, 14(2), 268-273.
- Giménez, A., Canto, J. M., Fernández, P., y Barrett, M. (2003). Stereotype development in andalusian children. *The Spanish Journal of Psychology*, 6(1), 28-34.
- Gomà-i-Freixanet, M., Grande, I., Valero i Ventura, S., y Puntí i Vidal, J. (2001). Personalidad y conducta delictiva autoinformada en adultos jóvenes. *Psicothema*, 13(2), 252-257.
- Gómez, J. A., Luengo, Á, y Romero, E. (2002a). Prevención del consumo de drogas en la escuela: Cuatro años de seguimiento de un programa. *Psicothema*, 14(4), 685-692.

- Gómez, G., y Acosta, M. V. (2002b). Valoración de la delgadez. un estudio transcultural (México/España). *Psicothema*, 14(2), 221-226.
- González, F., y García, V. (2000a). The role of ambiguous expectancy in differential inhibition: A different role for context from direct US association. *Psicológica*, 21(3), 213-232.
- González, J. A., Núñez, J. C., González, S., Álvarez, L., Roces, C., García, M., y col. (2000b). Autoconcepto, proceso de atribución causal y metas académicas en niños con y sin dificultades de aprendizaje. *Psicothema*, 12(4), 548-556.
- González, J. A., Núñez, J. C., Álvarez, L., González, S., Roces, C., González, P., y col. (2002). Inducción parental a la autorregulación, auto concepto y rendimiento académico. *Psicothema*, 14(4), 853-860.
- González, R. M. (2003). Propuesta de intervención en los procesos cognitivos y estructuras textuales en niños con DAE. *Psicothema*, 15(3), 458-463.
- Gorenkova, N. A., Nazarenko, I. V., Volkov, A. V., Avruschenko, M. S., Lapa, G. B., Kovalev, G. I., y col. (2005). Neuropsychological disorders indicative of postresuscitation encephalopathy in rats. *The Spanish Journal of Psychology*, 8(2), 246-255.
- Gracia, F. J., Arcos, J. L., y Caballer, A. (2000). Influencia de la presión temporal en el trabajo en grupo en función del tipo de tarea y del canal de comunicación. *Psicothema*, 12(2), 241-246.
- Gracia, E. (2002a). El maltrato infantil en el contexto de la conducta parental: Percepciones de padres e hijos. *Psicothema*, 14(2), 274-279.
- Gracia, F. J., Caballer, A., y Peiró, J. M. (2002b). Efectos de la presión temporal sobre la cohesión grupal en diferentes tipos de tareas y en diferentes canales de comunicación. *Psicothema*, 14(2), 434-439.
- Guerrero, C., y Palmero, F. (2006). Percepción de control y respuestas cardiovasculares. *International Journal of Clinical and Health Psychology*, 6(1), 145-168.
- Gutiérrez, J., y Arbej, J. (2005). Alexitimia y amplificación somatosensorial en el trastorno de pánico y en el trastorno de ansiedad generalizada. *Psicothema*, 17(1), 15-19.
- Guzmán, R., Jiménez, J. E., Ortiz, M. R., Hernández, I., Estévez, A., Rodrigo, M., y col. (2004). Evaluación de la velocidad de nombrar en las dificultades de aprendizaje de la lectura. *Psicothema*, 16(3), 442-447.
- Herrera, A., y Maldonado, A. (2002). Depresión, cognición y fracaso académico. *International Journal of Clinical and Health Psychology*, 2(1), 25-50.
- Herrero, Ó, Ordóñez, F., Salas, A., y Colom, R. (2002). Adolescencia y comportamiento antisocial. *Psicothema*, 14(2), 340-343.
- Herrero, M., y Viña, C. M. (2005). Conductas y actitudes hacia la alimentación en una muestra representativa de estudiantes de secundaria. *International Journal of Clinical and Health Psychology*, 5(1), 67-83.
- Higashiyama, A., Shimono, K., y Zaitzu, W. (2005). Contraction of perceived size and perceived depth in mirrors. *Psicológica*, 26(1), 81-95.
- Hubbard, T. L., Ruppel, S. E., y Courtney, J. R. (2005). The force of appearance: Gamma movement, naïve impetus, and representational momentum. *Psicológica*, 26(1), 209-228.
- Hughes, B. M. (2003). Self-esteem and changes in heart rate during laboratory-based stress. *Psicológica*, 24(1), 79-91.

- Hutchison, J. J., y Loomis, J. M. (2006). Does energy expenditure affect the perception of egocentric distance? A failure to replicate experiment 1 of proffitt, stefanucci, banton, and epstein (2003). *The Spanish Journal of Psychology*, 9(2), 332-339.
- Ibabe, I. (2000). Memoria de testigos: Recuerdo de acciones e información descriptiva de un suceso. *Psicothema*, 12(4), 574-578.
- Izura, C., y Ellis, A. W. (2002). Age of acquisition effects in word recognition and production in first and second languages. *Psicológica*, 23(2), 245-281.
- Krzemien, D., Urquijo, S., y Monchiatti, A. (2004). Aprendizaje social y estrategias de afrontamiento a los sucesos críticos del envejecimiento femenino. *Psicothema*, 16(3), 350-356.
- Labrador, F. J., Fernández-Alba, A., y Mañoso, V. (2002). Relación entre la reducción de las distorsiones cognitivas referidas al azar y la consecución de éxito terapéutico en jugadores patológicos. *Psicothema*, 14(3), 551-557.
- Labrador, F. J., Fernández, M. d. R., y Rincón, P. P. (2006). Eficacia de un programa de intervención individual y breve para el trastorno por estrés postraumático en mujeres víctimas de violencia doméstica. *International Journal of Clinical and Health Psychology*, 6(3), 527-547.
- Langley, L. K., Fuentes, L. J., Overmier, J. B., Bastin de Jong, C., y Prod'Homme, M. M. (2001). Attention to semantic and spatial information in aging and alzheimer's disease. *Psicológica*, 22(2), 293-323.
- Lillo, J., y Moreira, H. (2006). Lightness compression and hue changes. *The Spanish Journal of Psychology*, 9(2), 300-311.
- López, J. C., Vargas, J. P., Gómez, Y., y Salas, C. (2004a). Diferentes estrategias de aprendizaje espacial en tortugas reveladas mediante un procedimiento de inversión en un laberinto radial. *Psicológica*, 25(2), 147-162.
- López, M., Cantora, R., y Aguado, L. (2004b). Retardation and summation tests after extinction: The role of familiarity and generalization decrement. *Psicológica*, 25(1), 45-65.
- Lozano, L., y García, E. (2000). El rendimiento escolar y los trastornos emocionales y comportamentales. *Psicothema*, 12(Suppl2), 340-343.
- Luna, R., y Ruiz, M. (2001). Factores intrasujeto implicados en la percepción de la señalización vertical: Procesamiento holístico y analítico. *Psicothema*, 13(1), 141-146.
- Lupiáñez, J., Weaver, B., Tipper, S. P., y Madrid, E. (2001). The effects of practice on cueing in detection and discrimination tasks. *Psicológica*, 22(1), 1-23.
- Macizo, P., y Bajo, M. T. (2004a). Semantic facilitation and lexical competition in picture naming. *Psicológica*, 25(1), 1-22.
- Macizo, P., y Bajo, M. T. (2004b). When translation makes the difference: Sentence processing in reading and translation. *Psicológica*, 25(2), 181-205.
- Macizo, P., Bajo, T., y Soriano, M. F. (2006). Memoria operativa y control ejecutivo: Procesos inhibitorios en tareas de actualización y generación aleatoria. *Psicothema*, 18(1), 112-116.
- Manso, A. J., y Ballesteros, S. (2003). El papel de la agenda visoespacial en la adquisición del vocabulario ortográfico. *Psicothema*, 15(3), 388-394.
- Manteiga, R. D., y Chamizo, V. D. (2001). Aprendizaje elemental a pesar de entrenamiento configuracional en una tarea de navegación. *Psicológica*, 22(2), 235-252.

- Marcos, J. L., y Redondo, J. (2001). Relation between conditioned stimulus-elicited responses and unconditioned response diminution in long-interval human heart-rate classical conditioning. *The Spanish Journal of Psychology*, 4(1), 11-18.
- Marcos, J. L., y Redondo, J. (2002). Differential effects of expectancy and associative mechanisms on diminution of unconditioned response in electrodermal classical conditioning. *Psicothema*, 14(2), 375-381.
- Marcos, J. L., y Redondo, J. (2004). Modulación del reflejo de parpadeo mediante condicionamiento con imágenes aversivas como estímulos incondicionados. *Psicothema*, 16(3), 391-396.
- Marcos, J. L., y Redondo, J. (2005). Facilitation and interference of the automatic information processing on a reaction time task to threat-relevant stimuli. *Psicothema*, 17(2), 332-337.
- Marín, M., y León, J. M. (2001). Entrenamiento en habilidades sociales: Un método de enseñanza-aprendizaje para desarrollar las habilidades de comunicación interpersonal en el área de enfermería. *Psicothema*, 13(2), 247-251.
- Martín, I., de la Fuente, E. I., y Arnau, J. (2000). Manipulación del contenido en el fenómeno de ensombrecimiento en juicios de contingencia. *Psicothema*, 12(4), 599-604.
- Martín del Buey, F., y Camarero, F. (2001). Diferencias de género en los procesos de aprendizaje en universitarios.. *Psicothema*, 13(4), 598-604.
- Martínez, J., Muñiz, J., y García, E. (2000). Mejora de las puntuaciones de los tests mediante el entrenamiento. *Psicothema*, 12(Suppl2), 363-367.
- Martínez, J. M., y Robles, L. (2001). Variables de protección ante el consumo de alcohol y tabaco en adolescentes. *Psicothema*, 13(2), 222-228.
- Martínez, I. M., Cifre, E., Llorens, S., y Salanova, M. (2002). Efectos de la tecnología asistida por ordenador en el bienestar psicológico afectivo. *Psicothema*, 14(1), 118-123.
- Martínez, V., Zurriaga, R., Luque, O., y Moliner, C. (2005). Efecto modulador del tipo de segmento en la predicción de la satisfacción del consumidor. *Psicothema*, 17(2), 281-285.
- Marty, G. (2002). Formación de esquemas en el reconocimiento de estímulos estéticos. *Psicothema*, 14(1), 19-25.
- Marty, G., Munar, E., y Nadal, M. (2005). Familiaridad y evaluación de estímulos estéticos en función de la educación artística. *Psicothema*, 17(2), 338-343.
- Massironi, M., Rocchi, P., y Cornoldi, C. (2001). Does regularity affect the construction and memory of a mental image in the same way it affects a visual trace? *Psicológica*, 22(1), 115-142.
- Matsushima, E. H., de Oliveira, A. P., Ribeiro-Filho, N. P., y Da Silva, J. A. (2005). Visual angle as determinant factor for relative distance perception. *Psicológica*, 26(1), 97-104.
- Matud, M. P., García, M. A., y Matud, M. J. (2002). Estrés laboral y salud en el profesorado: Un análisis diferencial en función del género y del tipo de enseñanza. *International Journal of Clinical and Health Psychology*, 2(3), 451-465.
- Matud, M. P., Guerrero, K., y Matías, R. G. (2006). Relevancia de las variables sociodemográficas en las diferencias de género en depresión. *International Journal of Clinical and Health Psychology*, 6(1), 7-21.
- Mendez, X., Orgilés, M., y Espada, J. P. (2004). Escenificaciones emotivas para la fobia a la oscuridad: Un ensayo controlado. *International Journal of Clinical and Health Psychology*, 4(3), 505-520.

- Mendoza, E., Fresneda, M. D., Muñoz, J., Carballo, G., y Cruz, A. (2001). Morfología verbal: Estudio de las irregularizaciones de pseudoverbos en niños españoles. *Psicológica*, 22(2), 165-190.
- Miguel, M., y García, E. (2006). Influence of the typicality of the actions in a mugging script on retrieval-induced forgetting. *Psicológica*, 27(1), 119-135.
- Miñarro, J., Rodríguez, M., Pérez, N., Restall, C. J., y Brain, P. F. (2000). Developmental and neurochemical effects of early postnatal exposure to flumazenil in female and male mice. *Psicothema*, 12(4), 648-653.
- Molina, F. J., y Luciano, M. C. (2000). Complejidad de la tarea, demora de ejecución y comportamiento imitativo. *Psicothema*, 12(4), 561-567.
- Monsalve, A., y Cuetos, F. (2001). Asimetría hemisférica en el reconocimiento de palabras: Efectos de frecuencia e imaginabilidad. *Psicothema*, 13(1), 24-28.
- Moos, R. H., y Moos, B. S. (2006). Treated and untreated individuals with alcohol use disorders: Rates and predictors of remission and relapse. *International Journal of Clinical and Health Psychology*, 6(3), 513-526.
- Moreno, S., y Tudela, P. (2001). Visual attention and the reviewing process. *Psicothema*, 13(2), 277-283.
- Morera, M. D., Marichal, F. A., Quiles, M. N., Betancor, V., Rodríguez, R., Rodríguez, A., y col. (2004). La percepción de semejanza intergrupal y la identificación con el endogrupo: ¿incrementa o disminuye el prejuicio?. *Psicothema*, 16(1), 70-75.
- Moya, L., Salvador, A., Costa, R., Martínez, S., y González, E. (2003). Psychophysiological responses to acute stress in two groups of healthy women differing in fitness. *Psicothema*, 15(4), 563-568.
- Moya, L., Serrano, M. Á., González, E., Rodríguez, G., y Salvador, A. (2005). Respuesta psicofisiológica de estrés en una jornada laboral. *Psicothema*, 17(2), 205-211.
- Navarro, E., Reig, A., Barberá, E., y Ferrer, R. I. (2006). Grupo de iguales e iniciación sexual adolescente: Diferencias de género. *International Journal of Clinical and Health Psychology*, 6(1), 79-96.
- Navas, M., García, M. C., Rojas, A. J., Pumares, P., y Cuadrado, I. (2006). Actitudes de aculturación y prejuicio: La perspectiva de autóctonos e inmigrantes. *Psicothema*, 18(2), 187-193.
- Nievas, F., y Justicia, F. (2003). Development of memory structures for homographs using pathfinder network representations. *The Spanish Journal of Psychology*, 6(1), 12-27.
- Nievas, F., Justicia, F., Cañas, J. J., y Bajo, M. T. (2005). Lexical processing of ambiguous words: Dominance or associative strength? *The Spanish Journal of Psychology*, 8(2), 157-179.
- Novaes, M. E., Bignotto, M. M. B., Justo, A. P., y Gomes, T. M. (2006). Cardiovascular reactivity in hypertensives: Differential effect of expressing and inhibiting emotions during moments of interpersonal stress. *The Spanish Journal of Psychology*, 9(2), 154-161.
- Nuevo, R., y Montorio, I. (2005). Eficacia de un tratamiento cognitivo-conductual para la preocupación excesiva en personas mayores. *International Journal of Clinical and Health Psychology*, 5(2), 215-226.
- Olivares, J., y García, L. J. (2002). Resultados a largo plazo de un tratamiento en grupo para el miedo a hablar en público *Psicothema*, 14(2), 405-409.

- Olivares, J., Rosa, A. I., y Piqueras, J. A. (2005). Detección temprana y tratamiento de adolescentes con fobia social generalizada. *Psicothema*, 17(1), 1-8.
- Olivares, J., Rosa, A. I., y Olivares, P. J. (2006). Atención individualizada y tratamiento en grupo de adolescentes con fobia social generalizada. *International Journal of Clinical and Health Psychology*, 6(3), 565-580.
- Ortells, J. J., Vellido, C., Daza, M. T., y Noguera, C. (2006). Efectos de priming semántico con y sin conciencia perceptiva. *Psicológica*, 27(2), 225-242.
- Pacheco, V., Flores, C., González, F., Canales, C., y Carpio, C. (2005). Efectos de la consistencia e inconsistencia de las relaciones intrusivo-reforzador y muestra-reforzador en igualación a la muestra. *Psicothema*, 17(1), 118-122.
- Palmero, F., Breva, A., Diago, J. L., Díez, J. L., y García, I. (2002). Funcionamiento psicofisiológico y susceptibilidad a la sintomatología premenstrual en mujeres tipo A y tipo B. *International Journal of Clinical and Health Psychology*, 2(1), 111-136.
- Paterna, C., y Martínez, C. (2006). Fathers and gender traditionalism: Perception of inequality and life roles. *The Spanish Journal of Psychology*, 9(2), 171-181.
- Pellicer, O., Salvador, A., y Benet, I. A. (2002). Efectos de un estresor académico sobre las repuestas psicológica e inmune en jóvenes. *Psicothema*, 14(2), 317-322.
- Peñate, W., Perestelo, L., y Bethencourt, J. M. (2004). La predicción diferencial del nivel de depresión por las variables nivel de actividad, actitudes disfuncionales y estilo atributivo en función de la puntuación y la medida de depresión utilizada. *International Journal of Clinical and Health Psychology*, 4(1), 27-53.
- Perea, M. V., Ladera, V., y Morales, F. (2000). Aprendizaje verbal en el traumatismo craneoencefálico. *Psicothema*, 12(3), 353-359.
- Pérez, J. M., y Lubin, P. (2000). Estimación psicofísica de la "peligrosidad" del lanzamiento en balonmano. *Psicothema*, 12(Supl2), 438-441.
- Pérez, J. A., Lucas, A., Dasi, F., y Quiamzade, A. (2002a). La desobediencia masiva al código de circulación. normas heterónomas frente a normas inter-individuales. *Psicothema*, 14(4), 788-794.
- Pérez, A. M., Sanjuán, P., y Bermúdez, J. (2002b). Control personal y situacional y reactividad cardiovascular en tareas de estrés físico. *Psicothema*, 14(3), 583-590.
- Pérez, A., y Delgado, D. (2003). La codependencia en familias de consumidores y no consumidores de drogas: Estado del arte y construcción de un instrumento. *Psicothema*, 15(3), 381-387.
- Pérez, M. V., Díaz, A., y Vinet, E. (2005). Características psicológicas de adolescentes pertenecientes a comunidades educativas vulnerables. *Psicothema*, 17(1), 37-42.
- Philbeck, J. W., y O'Leary, S. (2005). Remembered landmarks enhance the precision of path integration. *Psicológica*, 26(1), 7-24.
- Pineño, O. (2004). Differential effects of nonreinforcement and punishment in humans. *Psicológica*, 25(1), 87-102.
- Pineño, O., y Matute, H. (2005). Outcome similarity modulates retroactive interference between cues trained apart. *Psicológica*, 26(2), 281-292.
- Pitarque, A., Algarabel, S., Dasí, C., y Ruiz, J. C. (2003). Olvido dirigido de falsas memorias: Podemos olvidar intencionadamente una falsa memoria?. *Psicothema*, 15(1), 6-11.
- Ponte, D., Sampedro, M. J., y Pardavila, M. (2004). Efecto de la excentricidad en tareas de búsqueda visual que difieren en las demandas atencionales. *Psicothema*, 16(4), 563-569.

- Prieto, M. D., López, O., Bermejo, M. R. B., Renzulli, J., y Castejón, J. L. (2002). Evaluación de un programa de desarrollo de la creatividad. *Psicothema*, 14(2), 410-414.
- Quero, S., Baños, R. M., y Botella, C. (2000). Sesgos cognitivos en el trastorno de pánico: Comparación entre el stroop computerizado y con tarjetas. *Psicothema*, 12(2), 165-170.
- Quiles, M. N., Leyens, J. P., y Rodríguez, A. (2001). Atribuciones a situaciones de interacción ambiguas: El punto de vista de la víctima y del agente del prejuicio. *Psicothema*, 13(4), 557-562.
- Quiles, M. N., Leyens, J. P., Rodríguez, A., Betancor, V., y Rodríguez, R. (2003). El género como criterio de pertenencia grupal y las atribuciones al prejuicio. *Psicothema*, 15(1), 65-70.
- Ramos, M. M., y Catena, A. (2005). The dissociation between the recall of stimulus frequencies and the judgment of contingency allows the placement of the competition effect in the final causal processing stages. *Psicológica*, 26(2), 293-303.
- Rebollo, I., Herrero, Ó, y Colom, R. (2002). Personality in imprisoned and non-imprisoned people: Evidence from the EPQ-R. *Psicothema*, 14(3), 540-543.
- Recarte, M. Á, Conchillo, Á, y Nunes, L. M. (2005). Estimation of arrival time in vehicle and video. *Psicothema*, 17(1), 112-117.
- Rial, A., Varela, J., y García, A. (2001). Identificación de las variables que determinan la evaluación de una extensión de marca. *Psicothema*, 13(2), 284-289.
- Rodrigo, T., Arall, M., y Chamizo, V. D. (2005). Blocking and unblocking in a navigation task. *Psicológica*, 26(2), 229-241.
- Rodríguez, R., y Rodríguez, A. (2000). El efecto de la supresión de pensamientos estereotípicos en una tarea cognitiva. *Psicothema*, 12(1), 33-39.
- Rodríguez, G., y Alonso, G. (2002). Efecto de las intensidad del estímulo en la inhibición latente: Un caso de decremento en la generalización. *Psicológica*, 23(2), 401-413.
- Rodríguez, R., y Moya, M. (2003). La diferencia percibida en valores como proceso vinculado a las relaciones intergrupales de payos y gitanos. *Psicothema*, 15(2), 176-182.
- Rodríguez, G., y Alonso, G. (2004). Latent inhibition as a function of US intensity in a two-stage CER procedure. *Psicológica*, 25(2), 207-216.
- Rodríguez, A., Betancor, V., Rodríguez, R., Quiles, M. N., Delgado, N., y Coello, E. (2005a). El efecto de las identidades nacionales con distintos niveles de inclusividad en el prejuicio hacia exogrupos. *Psicothema*, 17(3), 441-446.
- Rodriguez, G., Triana, B., y García, Hernández, M. (2005b). La experiencia familiar y la atribución de roles parentales. *Psicothema*, 17(3), 363-369.
- Rodríguez, R., Moya, M., y Yzerbyt, V. (2006a). Cuando el poder ostentado es innmerecido: Sus efectos sobre la percepción y los juicios sociales. *Psicothema*, 18(2), 194-199.
- Rodríguez, A., Coello, E., Betancor, V., Rodríguez, R., y Delgado, N. (2006b). Amenaza al endogrupo y nivel de infrahumanización del exogrupo. *Psicothema*, 18(1), 73-77.
- Rodríguez, G., Camacho, J., Rodrigo, M. J., Martín, J. C., y Máiquez, M. L. (2006c). Evaluación del riesgo psicosocial en familias usuarias de servicios sociales municipales. *Psicothema*, 18(2), 200-206.
- Romero, M. A., Vila, N. J., y Rosas, J. M. (2003). Time and context effects after discrimination reversal in human beings. *Psicológica*, 24(2), 169-184.



- Rosas, J. M., García, A., y Callejas, J. E. (2006). Effects of context change upon retrieval of first and second-learned information in human predictive learning. *Psicológica*, 27(1), 35-56.
- Ruipérez, M. A., García-Palacios, A., y Botella, C. (2002). Clinical features and treatment response in social phobia: Axis II comorbidity and social phobia subtypes. *Psicothema*, 14(2), 426-433.
- Ruiz, J. A., y Sánchez, C. (2001). Depresión y memoria: Es la información congruente con el estado de ánimo más accesible?. *Psicothema*, 13(2), 193-196.
- Ruiz, E., Sánchez, M. B., Thayer, J. F., y Vila, J. (2002). Modulación no consciente de la respuesta cardiaca de defensa por imágenes fóbicas. *Psicothema*, 14(4), 739-745.
- Ruiz, I., de Labry, A. O., Delgado, C. J., Marcos, M., Muñoz, N., Pascuau, J., y col. (2005). Impacto del apoyo social y la morbilidad psíquica en la calidad de vida en pacientes tratados con antirretrovirales. *Psicothema*, 17(2), 245-249.
- Sánchez, M., Ruiz, E., Pérez, N., Fernández, M. C., Cobos, P., y Vila, J. (2002). Modulación emocional de los reflejos defensivos mediante visualización de imágenes afectivas. *Psicothema*, 14(4), 702-707.
- Sánchez, J. P., Martínez, J. M., Román, F., y Torrente, G. (2006a). The effect of content and physical properties of affective pictures on emotional responses. *The Spanish Journal of Psychology*, 9(2), 145-153.
- Sánchez, J. C. (2006b). Efectos de la presentación del mensaje para realizar conductas saludables: El papel de la autoeficacia y de la motivación cognitiva. *International Journal of Clinical and Health Psychology*, 6(3), 613-630.
- Sansa, J., y Prados, J. (2003). Ensombrecimiento entre puntos de referenda en una tarea de navegación. *Psicológica*, 24(1), 17-29.
- Santamaría, C., y Espino, O. (2006). Pensar en lo verdadero para seleccionar lo falso. *Psicológica*, 27(2), 195-206.
- Santostefano, S., Quiroga, M. A., y Rooney, S. (2001). Life stressors and cognitive styles in children. *The Spanish Journal of Psychology*, 4(1), 37-47.
- Sanz de Acedo, M. L. y Iriarte, M. D. (2001). Enhancement of cognitive functioning and self-regulation of learning in adolescents. *The Spanish Journal of Psychology*, 4(1), 55-64.
- Sebastián, M. V., y Elosúa, M. R. (2002). Tipo de material y olvido a corto plazo en pacientes alzheimer y ancianos sanos. *Psicothema*, 14(2), 199-204.
- Sebastian, A., Rodríguez, G., y Alonso, G. (2006). El efecto espinet en aversión condicionada al sabor: Resultados negativos. *Psicológica*, 27(1), 97-117.
- Sierra, B., Falces, C., y Briñol, P. (2002). Recuerdo sobre situaciones reales basadas en guiones: Relevancia y tipicidad de las acciones. *Psicothema*, 14(4), 776-782.
- Sierra, J. C., y Buela, G. (2003). Efectos de una dosis aguda de diacepam sobre la expectación en sujetos voluntarios: Diferencias de género. *Psicothema*, 15(2), 211-215.
- Singh-Manoux, A. (2000). Culture and gender issues in adolescence: Evidence from studies on emotion. *Psicothema*, 12(Suppl), 93-100.
- Sipsma, E., Carroble, J., Montorio, I., y Everaerd, W. (2000). Sexual aggression against women by men acquaintances: Attitudes and experiences among spanish university students. *The Spanish Journal of Psychology*, 3(1), 14-27.

- 
- Soares, A. P., y de Sousa, C. M. (2006). Estilo de vida, percepção de saúde e estado de saúde em estudantes universitários portugueses: Influência da área de formação. *International Journal of Clinical and Health Psychology*, 6(3), 647-663.
- Soriano, M. F., Macizo, P., y Bajo, T. (2004). Diferencias individuadas en tareas de interferencia episódica y semántica. *Psicothema*, 16(2), 187-193.
- Soto, S., y Sebastián, N. (2002). Identity and similarity in repetition deafness. *Psicológica*, 23(2), 189-208.
- Tornay, F. J., y Milán, E. G. (2001). Diferente participación de los mecanismos de control en el cambio entre tareas regular frente al cambio al azar. *Psicothema*, 13(1), 111-117.
- Toro, J., y Varas, N. (2004). Los otros: Prejuicio y distancia social hacia hombres gay y lesbianas en una muestra de estudiantes de nivel universitario. *International Journal of Clinical and Health Psychology*, 4(3), 537-551.
- Tortella, M., Servera, M., Balle, M., y Fullana, M. A. (2004). Viabilidad de un programa de prevención selectiva de los problemas de ansiedad en la infancia aplicado en la escuela. *International Journal of Clinical and Health Psychology*, 4(2), 371-387.
- Tous, J. M., Viadé, A., y Chico, E. (2003). Aplicación del psicodiagnóstico miokinético revisado (PMK-R) al estudio de la violencia. *Psicothema*, 15(2), 253-259.
- Trujano, P., y Raich I Escursell, M. (2000). Variables socioculturales en la atribución de culpa a las víctimas de violación. *Psicothema*, 12(2), 223-228.
- Trujillo, H. M., Oviedo, E., y Vargas, C. (2006). Compensatory and mimetic conditioned responses to effects of heroin in addicted persons. *Psicothema*, 18(1), 59-65.
- Valiente, R. M., Sandín, B., Chorot, P., y Tabar, A. (2003). Diferencias según la edad en la prevalencia e intensidad de los miedos durante la infancia y la adolescencia: Datos basados en el FSSC-R. *Psicothema*, 15(3), 414-419.
- Vázquez, C., Nieto, M., Cerviño, M. J., y Fuentenebro, F. (2006). Efectos del incremento de la demanda cognitiva en tareas de atención sostenida en los trastornos esquizofrénicos y la esquizotipia. *Psicothema*, 18(2), 221-227.
- Vega, Z., Vila, J., y Rosas, J. M. (2004). Efecto de devaluación e inflación de la consecuencia sobre la asociación respuesta-consecuencia en una tarea instrumental con humanos. *Psicológica*, 25(2), 163-179.
- Vera, P. E., Pérez, V., Moreno, E., y Allende, F. (2004). Diferencias en variables psicosociales en sujetos VIH homosexuales y heterosexuales. *International Journal of Clinical and Health Psychology*, 4(1), 55-67.
- Verdugo, M. Á., y Sabeh, E. N. (2002). Evaluación de la percepción de calidad de vida en la infancia. *Psicothema*, 14(1), 86-91.
- Vigil, A., Pérez, J., y García, J. E. (2000). The role of the P300 component in a translation-recognition task. *Psicothema*, 12(4), 605-614.
- Vinader, C., Ferrer, A., Arenas, M. C., Monleón, S., y Parra, A. (2002). La maprotilina anula las diferencias entre ratones machos y hembras en el laberinto de agua de morris. *Psicothema*, 14(4), 823-827.
- Viñas, F., Canals, J., Gras, M. E., Ros, C., y Domènech, E. (2002). Psychological and family factors associated with suicidal ideation in pre-adolescents. *The Spanish Journal of Psychology*, 5(1), 20-28.
- Vizcarral, M. B., Balladares, E., Candia, C., Lepe, M., y Saldivia, C. (2004). Conducta sexual durante la infancia en estudiantes chilenos. *Psicothema*, 16(1), 58-63.

Yela, J. R., Gómez, M. Á, y Salgado, A. (2001). Efectos de la predecibilidad en indefensión aprendida crónica. *Psicothema*, 13(4), 592-597.

## ANEXO 2. Valoraciones de los artículos revisados

SJP: The Spanish Journal of Psychology

IJCHP: International Journal of Clinical and Health Psychology

### **Claves utilizadas para la clasificación de los artículos según la forma de interpretar una interacción significativa**

- 2 Factores:    A: no analiza / análisis incompleto / no interpreta  
                  B: interpreta interacción por medio de los efectos simples  
                  C: análisis e interpretación correctas
- 3 Factores:    A: no analiza / análisis incompleto / no interpreta  
                  B1: interpreta interacción por medio de los efectos simples simples  
                  B2: interpreta interacción por medio de los efectos simples  
                  B3: interpreta interacción por medio de los efectos simples de interacción  
                  C: análisis e interpretación correctas

nº	Revista	Año	Volumen	Páginas	Clasificación	
					2 Factores	3 Factores
1	Psicothema	2003	15(4)	589-594	B	
2	SJP	2006	9(2)	312-320	B	
3	IJCHP	2003	3(1)	9-21	B	
4	Psicológica	2005	26(2)	261-279	B	
5	Psicológica	2005	26(1)	61-79	C	
6	Psicothema	2000	12(supl2)	15-17	C	
7	Psicothema	2006	18(1)	143-148	B	
8	Psicothema	2000	12(supl2)	22-24	B	
9	Psicothema	2001	13(2)	302-309	B	
10	Psicothema	2002	14(2)	393-398	B	
11	Psicológica	2006	27(1)	57-77	C	B2, C
12	Psicothema	2005	17(1)	76-82	A	
13	Psicothema	2004	16(2)	276-281	A	A
14	Psicothema	2002	14(3)	597-604	A, B	
15	Psicothema	2004	16(1)	50-57	B	A
16	Psicothema	2005	17(1)	71-75	C	
17	Psicothema	2002	14(1)	47-52	B	
18	Psicothema	2002	14(2)	209-214	B	A
19	Psicothema	2004	16(1)	1-6	B	
20	Psicothema	2002	14(4)	746-753	A	
21	SJP	2006	9(1)	10-18	B	
22	Psicothema	2000	12(1)	71-78	B	
23	IJCHP	2002	2(3)	439-450	B	A
24	Psicológica	2003	24(1)	31-55	A, B	
25	Psicológica	2001	22(2)	217-234	B	
26	Psicothema	2006	18(1)	88-99	B	
27	Psicothema	2000	12(supl2)	60-63	B	
28	IJCHP	2005	5(1)	23-42	B, C	
29	Psicothema	2003	15(3)	401-406	A	
30	Psicothema	2004	16(1)	14-21	B	
31	Psicothema	2005	17(4)	559-562	A	
32	Psicothema	2005	17(1)	64-70	B	
33	Psicothema	2005	17(3)	370-374	B	
34	Psicothema	2000	12(4)	557-560	B	
35	Psicothema	2001	13(2)	318-323	B	
36	Psicothema	2003	15(3)	407-413	C	C
37	Psicothema	2005	17(3)	447-452	B	
38	Psicothema	2005	17(2)	297-302	B	B2
39	Psicológica	2001	22(2)	267-291	B	
40	Psicológica	2002	23(2)	371-400	B	
41	SJP	2004	7(1)	63-68	B	
42	Psicológica	2001	22(1)	97-113	C	A
43	Psicothema	2005	17(4)	590-596	B	
44	Psicothema	2000	12(4)	586-593	B	
45	Psicothema	2002	14(4)	771-775	B	
46	Psicothema	2004	16(4)	606-610	B	
47	Psicothema	2004	16(1)	27-31	B, C	
48	Psicothema	2005	17(3)	396-403	B	
49	Psicológica	2001	22(2)	205-215	A, B	

n°	Revista	Año	Volumen	Páginas	Clasificación	
					2 Factores	3 Factores
50	SJP	2006	9(2)	201-218	B	A
51	IJCHP	2006	6(2)	271-284	B	
52	Psicológica	2003	24(2)	185-200	B	
53	SJP	2005	8(1)	3-11	A, B	A
54	Psicothema	2000	12(4)	615-622	B	B2
55	Psicothema	2000	12(supl2)	108-110	A	
56	Psicológica	2006	27(2)	243-267	B	
57	Psicothema	2000	12(3)	360-367	B	
58	Psicothema	2004	16(2)	229-234	B	
59	Psicothema	2001	13(1)	118-126	A	
60	Psicológica	2004	25(1)	67-85	B	
61	Psicothema	2003	15(4)	631-637	B	A
62	Psicothema	2004	16(1)	104-109	B	
63	Psicothema	2004	16(2)	248-254	B	
64	Psicothema	2000	12(2)	229-235	A, B	
65	Psicothema	2004	16(4)	611-614	B	
66	Psicothema	2002	14(1)	106-111	B	
67	Psicothema	2000	12(supl2)	152-156	A	
68	SJP	2006	9(1)	32-37	A	
69	Psicothema	2000	12(3)	426-434	A, B	
70	Psicothema	2003	15(2)	216-220	B	A
71	Psicothema	2005	17(4)	563-568	A	
72	Psicothema	2004	16(2)	270-275	B	
73	Psicothema	2002	14(2)	293-299	B	
74	SJP	2006	9(2)	285-294	A	
75	Psicothema	2000	12(supl2)	163-165	B	
76	Psicológica	2000	21(1-2)	23-37	B	
77	Psicothema	2000	12(2)	285-292	B	
78	Psicothema	2002	14(1)	53-62	B	
79	Psicothema	2000	12(2)	180-186	A, B	
80	Psicothema	2006	18(1)	1-8	B	
81	Psicothema	2004	16(4)	654-660	B	
82	SJP	2005	8(1)	68-78	B	
83	Psicológica	2006	27(1)	79-95	A	A
84	Psicothema	2003	15(2)	234-240	B	
85	Psicológica	2004	25(2)	135-146	B	
86	Psicothema	2002	14(3)	572-576	B	
87	Psicothema	2000	12(4)	520-524	A	
88	Psicothema	2000	12(2)	201-206	B	
89	Psicológica	2003	24(2)	159-167	B	
90	Psicothema	2002	14(2)	357-362	B	
91	Psicothema	2004	16(2)	181-186	B	
92	Psicothema	2002	14(4)	783-787	B	A
93	Psicothema	2001	13(4)	622-628	B	
94	Psicológica	2001	22(2)	253-266	C	
95	Psicothema	2002	14(4)	795-801	B	
96	IJCHP	2004	4(3)	481-493	A	
97	Psicothema	2004	16(2)	235-240	B	
98	Psicothema	2005	17(4)	607-613	B	

n°	Revista	Año	Volumen	Páginas	Clasificación	
					2 Factores	3 Factores
99	Psicothema	2000	12(supl2)	211-213	A	A
100	SJP	2002	5(2)	78-89	A	
101	IJCHP	2006	6(1)	97-125	B	
102	Psicothema	2003	15(3)	369-374	B	
103	Psicológica	2005	26	243-260	B	
104	SJP	2006	9(2)	321-331	A	
105	SJP	2001	4(1)	19-25	B, C	
106	Psicothema	2003	15(2)	260-266	C	
107	SJP	2005	8(1)	12-20	B, C	A
108	Psicológica	2005	26(1)	147-160	A	A
109	Psicothema	2004	16(3)	429-435	B	
110	SJP	2000	3(1)	28-36	B	
111	Psicothema	2000	12(2)	207-211	B	
112	Psicothema	2001	13(4)	611-616	A	A
113	Psicothema	2002	14(2)	450-455	B	B1, B3
114	Psicológica	2003	24(2)	271-287	B	
115	Psicológica	2003	24(2)	243-269	B	
116	Psicothema	2004	16(2)	194-202	C	
117	Psicothema	2004	16(2)	211-216	B	B1
118	IJCHP	2005	5(2)	385-392	B	
119	Psicothema	2005	17(1)	49-56	B	
120	Psicothema	2004	16(2)	294-298	B	
121	Psicológica	2005	26(1)	121-137	C	
122	Psicothema	2002	14(2)	268-273	A, B	
123	SJP	2003	6(1)	28-34	B	
124	Psicothema	2001	13(2)	252-257	B	
125	Psicothema	2002	14(4)	685-692	B	
126	Psicothema	2002	14(2)	221-226	B	
127	Psicológica	2000	21	213-232	B	A, B2, B3
128	Psicothema	2000	12(4)	548-556	B	B2
129	Psicothema	2002	14(4)	853-860	B	B2
130	Psicothema	2003	15(3)	458-463	A, B	
131	SJP	2005	8(2)	246-255	B	
132	Psicothema	2000	12(2)	241-246	A	
133	Psicothema	2002	14(2)	274-279	B	
134	Psicothema	2002	14(2)	434-439	B	
135	IJCHP	2006	6(1)	145-168	B	
136	Psicothema	2005	17(1)	15-19	B	
137	Psicothema	2004	16(3)	442-447	B	
138	IJCHP	2002	2(1)	25-50	B	
139	Psicothema	2002	14(2)	340-343	B	
140	IJCHP	2005	5(1)	67-83	A	
141	Psicológica	2005	26(1)	81-95	A, B	
142	Psicológica	2005	26(1)	209-228	A	
143	Psicológica	2003	24(1)	79-91	B	
144	SJP	2006	9(2)	332-339	B	
145	Psicothema	2000	12(4)	574-578	B	
146	Psicológica	2002	23(2)	245-281	B	
147	Psicothema	2004	16(3)	350-356	B	

n°	Revista	Año	Volumen	Páginas	Clasificación	
					2 Factores	3 Factores
148	Psicothema	2002	14(3)	551-557	B	A
149	IJCHP	2006	6(3)	527-547	B	
150	Psicológica	2001	22(2)	292-323	A	
151	Psicothema	2001	13(1)	159-165	B	
152	SJP	2006	9(2)	300-311	B	
153	Psicológica	2004	25(1)	45-65	B	
154	Psicológica	2004	25(2)	147-162	B	
155	Psicothema	2000	12(supl2)	340-343	B	
156	Psicothema	2001	13(1)	141-146	B	
157	Psicológica	2001	22(1)	1-23	C	A, B2
158	Psicológica	2004	25(1)	1-22	B	
159	Psicológica	2004	25(2)	181-205	B	
160	Psicothema	2006	18(1)	112-116	B	
161	Psicothema	2003	15(3)	388-394	B, C	
162	Psicológica	2001	22(2)	235-252	B	B3
163	SJP	2001	4(1)	11-18	A	
164	Psicothema	2002	14(2)	375-381	B	
165	Psicothema	2004	16(3)	391-396	B, C	A
166	Psicothema	2005	17(2)	332-337	B	B1
167	Psicothema	2001	13(2)	247-251	A, B	
168	Psicothema	2000	12(4)	599-604	B	
169	Psicothema	2001	13(4)	598-604	B	
170	Psicothema	2000	12(supl2)	363-367	B	
171	Psicothema	2001	13(2)	222-228	A	
172	Psicothema	2002	14(1)	118-123	B	
173	Psicothema	2005	17(2)	281-285	C	
174	Psicothema	2002	14(1)	19-25	C	
175	Psicothema	2005	17(2)	338-343	B	A
176	Psicológica	2001	22(1)	115-142	A, B	A
177	Psicológica	2005	26(1)	97-104	A	A
178	IJCHP	2002	2(3)	451-465	B	
179	IJCHP	2006	6(1)	7-21	B	
180	IJCHP	2004	4(3)	505-520	A, B	
181	Psicológica	2001	22(2)	165-190	B	
182	Psicológica	2006	27(1)	119-135	B	
183	Psicothema	2000	12(4)	648-653	B	
184	Psicothema	2000	12(4)	561-567	A	A
185	Psicothema	2001	13(1)	24-28	A, B	A
186	IJCHP	2006	6(3)	513-526	A	
187	Psicothema	2001	13(2)	277-283	B	
188	Psicothema	2004	16(1)	70-75	B	
189	Psicothema	2003	15(4)	563-568	B	
190	Psicothema	2005	17(2)	205-211	B	
191	IJCHP	2006	6(1)	79-96	B	B1
192	Psicothema	2006	18(2)	187-193	B	B
193	SJP	2003	6(1)	12-27	B	A
194	SJP	2005	8(2)	157-179	A, B	
195	SJP	2006	9(2)	154-161	B	A
196	IJCHP	2005	5(2)	215-226	B	



n°	Revista	Año	Volumen	Páginas	Clasificación	
					2 Factores	3 Factores
197	Psicothema	2002	14(2)	405-409	B	
198	Psicothema	2005	17(1)	1-8	B	
199	IJCHP	2006	6(3)	565-580	B	
200	Psicológica	2006	27(2)	225-242	B	
201	Psicothema	2005	17(1)	118-122	A	
202	IJCHP	2002	2(1)	111-136	B	
203	SJP	2006	9(2)	171-181	B	B1
204	Psicothema	2002	14(2)	317-322	B	
205	IJCHP	2005	4(1)	27-53	B	
206	Psicothema	2000	12(3)	353-359	B	
207	Psicothema	2000	12(supl2)	438-441	B	A
208	Psicothema	2002	14(4)	788-794	C	
209	Psicothema	2002	14(3)	583-590	B	A
210	Psicothema	2003	15(3)	381-387	B	A
211	Psicothema	2005	17(1)	37-42	B	A
212	Psicológica	2005	26(1)	7-24	B	
213	Psicológica	2004	25(1)	87-102	B	
214	Psicológica	2005	26(2)	281-292	A, B	
215	Psicothema	2003	15(1)	6-11	C	
216	Psicothema	2004	16(4)	563-569	A	
217	Psicothema	2002	14(2)	410-414	C	A
218	Psicothema	2000	12(2)	165-170	C	
219	Psicothema	2001	13(4)	557-562	B	
220	Psicothema	2003	15(1)	65-70	B	
221	Psicológica	2005	26(2)	293-303	B	
222	Psicothema	2002	14(3)	540-543	B	A
223	Psicothema	2005	17(1)	112-117	A	
224	Psicothema	2001	13(2)	284-289	B	
225	Psicológica	2005	26(2)	229-241	B	
226	Psicothema	2000	12(1)	33-39		C
227	Psicológica	2002	23(2)	401-413	B	
228	Psicothema	2003	15(2)	176-182	B	B1
229	Psicológica	2004	25(2)	207-216	B	B1, B3
230	Psicothema	2005	17(3)	441-446	B	A
231	Psicothema	2005	17(3)	363-369		A
232	Psicothema	2006	18(2)	194-199	B	
233	Psicothema	2006	18(1)	73-77	B	B2
234	Psicothema	2006	18(2)	200-206	B	
235	Psicológica	2003	24(2)	169-84	B	
236	Psicológica	2006	27(1)	35-56	B	A
237	Psicothema	2001	13(4)	426-433	B	A
238	Psicothema	2001	13(2)	193-196	B	
239	Psicothema	2002	14(4)	739-745	B	
240	Psicothema	2005	17(2)	245-249	B	
241	Psicothema	2002	14(4)	702-707	B	B1, B2
242	SJP	2006	9(2)	145-153	B	
243	IJCHP	2006	6(3)	613-630	B	B3
244	Psicológica	2003	24(1)	17-29	A, B	
245	Psicológica	2006	27(2)	195-206	B	

n°	Revista	Año	Vololumen	Páginas	Clasificación	
					2 Factores	3 Factores
246	SJP	2001	4(1)	37-47	B	
247	SJP	2001	4(1)	55-64	B	
248	Psicothema	2002	14(2)	199-204	B	
249	Psicológica	2006	27(1)	97-117	B	
250	Psicothema	2002	14(4)	776-782	B, C	
251	Psicothema	2003	15(2)	211-215	B	
252	Psicothema	2000	12(supl1)	93-100	B	
253	SJP	2000	3(1)	14-27	A	A
254	IJCHP	2006	6(3)	647-663	B	
255	Psicothema	2004	16(2)	187-193	B	A
256	Psicológica	2002	23(2)	189-208	B	
257	Psicothema	2001	13(1)	111-117	B	B3
258	IJCHP	2004	4(3)	537-551	A	A
259	IJCHP	2004	4(2)	371-387	B	
260	Psicothema	2003	15(2)	253-259	A, B	
261	Psicothema	2000	12(2)	223-228	A	
262	Psicothema	2006	18(1)	59-65		B2
263	Psicothema	2003	15(3)	414-419	A	
264	Psicothema	2006	18(2)	221-227	B	
265	Psicológica	2004	25(2)	163-179	B	
266	IJCHP	2004	4(1)	55-67	B	
267	Psicothema	2002	14(1)	86-91	A	A
268	Psicothema	2000	12(4)	605-614	C	A
269	Psicothema	2002	14(4)	823-827	B	A
270	SJP	2002	5(1)	20-28	B	
271	Psicothema	2004	16(1)	58-63	B	
272	Psicothema	2001	13(4)	592-597	A, B	

## ANEXO 3. Índice de tablas

Tabla	Título	Página
2.1	Notación utilizada en un diseño $2 \times 3$	14
2.2	Notación utilizada en un diseño $2 \times 2 \times 2$	18
3.1	Clasificación de las interacciones entre dos factores en función del análisis e interpretación realizados por los investigadores	26
3.2	Datos de un diseño $2 \times 2$	30
3.3	Medias correspondientes a los datos de la Tabla 3.2	31
3.4	Tabla de ANOVA	31
3.5	Análisis de los efectos simples	31
3.6	Datos de un diseño $2 \times 3$	32
3.7	Medias correspondientes a los datos de la Tabla 3.6	32
3.8	Tabla de ANOVA	33
3.9	Análisis de los efectos simples	33
3.10	Clasificación de las interacciones entre tres factores en función del análisis e interpretación realizados por los investigadores	39
3.11	Notación en un diseño $2 \times 2 \times 3$	41
3.12	Notación en un diseño $3 \times 2 \times 2$ en el que se ha agrupado el tercer factor (C), celdas en fondo gris, para presentar la notación correspondiente	43
3.13	Estadísticos descriptivos (tomado de Howell, 2002, pág. 432)	48
3.14	Tabla de ANOVA (tomado de Howell, 2002, pág. 432)	48
3.15	Efectos simples (tomado de Howell, 2002, pág. 432)	48
3.16	Recomendaciones de los manuales de análisis de datos con SPSS para interpretar la interacción significativa	55
4.1	Tomada de Verdugo y Sabeh (2002, pág. 89)	62
4.2	Tomado de Cardenal y Fierro (2001, pág. 121)	64
4.3	Tomado de Cardenal y Fierro (2001, pág. 122)	64
4.4	Tomado de Amigo y col. (2002, pág. 49)	68
4.5	Tomado de Gonzalez (2003, pág. 461)	80
4.6	Tomada de Nuevo y Montorio (2005, pág. 220)	84
4.7	Tomada de Ruiz y Sánchez (2001, pág. 194)	101
4.8	Tomado de Martínez y col. (2000; pág. 366)	105
4.9	Tomado de Soriano, Macizo y Bajo (2004, pág. 191)	108
4.10	Tomado de Prieto y col. (2002, pág. 412)	112
4.11	Tomado de García y Arias-Gundín (2004, pág. 199)	113
4.12	Tomado de Pitarque y col. (2003, pág. 8)	115
4.13	Tomada de Rodríguez y Moya (2003, pág. 179)	123
4.14	Tomada de Marcos y Redondo (2005, pág. 335)	133
4.15	Tomada de Marcos y Redondo (2005, pág. 335)	133
4.16	Tomado de Algarabel, Luciano y Martínez (2006, pág. 65)	144
5.1	Notación utilizada en un diseño $2 \times 3$	148
5.2	Coefficientes para comparaciones lineales en un diseño $2 \times 3$	151
5.3	Notación utilizada en un diseño $3 \times 3$	153
5.4	Coefficientes para comparar efectos simples en un diseño $3 \times 3$	154

Tabla	Título	Página
5.5	Notación utilizada en un diseño 3×4	155
5.6	Coeficientes para comparar efectos simples en un diseño 3×4	157
5.7	Diseño 2×3. Datos del ejemplo 2 con la nueva variable <i>C</i>	162
5.8	Resultados de las comparaciones $\psi_1$ , $\psi_2$ y $\psi_3$	163
5.9	Diseño 2×3. Datos del ejemplo 2 modificados	164
5.10	Medias del diseño 2×3 propuesto en la Tabla 5.9	165
5.11	Tabla resumen del ANOVA	165
5.12	Significación de los efectos simples de <i>A</i>	165
5.13	Resultados de las comparaciones $\psi_1$ , $\psi_2$ y $\psi_3$	166
5.14	Algunas comparaciones planeadas para interpretar la interacción del ejemplo presentado por Howell (2002, pág. 425)	167
5.15	Notación utilizada en un diseño 2×5	168
5.16	Efecto de la <i>edad</i> : diferencia entre las dos tareas de bajo procesamiento y las tres de alto procesamiento (resultados obtenidos a partir de los datos presentados por Howell, 2002, pág. 425)	169
5.17	Resultados de las comparaciones $\psi_1$ , $\psi_2$ y $\psi_3$ con LMATRIX	173
5.18	Efecto global de las comparaciones $\psi_1$ , $\psi_2$ y $\psi_3$	173
5.19	Notación utilizada en un diseño 2×2×3	178
5.20	Coeficientes para efectuar comparaciones lineales en un diseño 2×2×2	179
5.21	Notación utilizada en un diseño 2×3×3	183
5.22	Coeficientes para efectuar comparaciones entre efectos simples de interacción en un diseño 2×3×3	187
5.23	Notación utilizada en un diseño 3×3×3	189
5.24	Coeficientes para efectuar comparaciones entre efectos simples de interacción en un diseño 3×3×3	192
5.25	Datos del ejemplo	195
5.26	Medias observadas en cada combinación de niveles	196
5.27	Tabla resumen del ANOVA	196
5.28	Comparaciones entre los efectos simples de interacción	197
5.29	Coeficientes para valorar la significación individual de los efectos simples asociados a la interacción entre <i>tratamiento</i> y <i>hacinamiento</i>	200
5.30	Significación individual de los efectos simples de interacción	200
5.31	Significación de la interacción entre <i>tratamiento</i> y <i>hacinamiento</i> en cada <i>momento del día</i> (significación individual de los efectos simples de interacción)	201
5.32	Resultados de las comparaciones $\psi_1$ , $\psi_2$ y $\psi_3$ con LMATRIX	205
5.33	Efecto global de las comparaciones $\psi_1$ , $\psi_2$ y $\psi_3$	205

## ANEXO 4. Índice de figuras

Figura	Título	Página
3.1	Distintas pautas de interacción en un diseño 2x2	35
3.2	Tomado de Howell (2002, pág. 431)	49
3.3	Tomado de Maxwell y Delanay (2004, pág. 308)	50
4.1	Tomada de Alonso y Román (2005, pág. 79)	58
4.2	Tomada de Chico (2000, pág. 231)	60
4.3	Tomada de Fernández y col. (2004, pág. 488)	61
4.4	Tomada de Vera y col. (2004, pág. 62)	66
4.5	Tomada de Vera y col. (2004, pág. 62)	67
4.6	Tomado de Pacheco y otros (2005, pág. 120)	71
4.7	Tomado de Durán y col. (2000, pág. 523)	72
4.8	Tomado de Gracia y col. (2000, pág. 244)	74
4.9	Tomado de Gracia y col. (2000, pág. 244)	75
4.10	Tomado de Alvarado y Santisteban (2004, pág. 279)	77
4.11	Tomado de Monsalve y Cuetos (2001, pág. 26)	79
4.12	Representación gráfica a partir de la Tabla 2 de González (2003; pág. 461)	82
4.13	Tomada de Rodríguez y col. (2005, pág. 244)	87
4.14	Tomado de Algarabel, Pitarque y Dasí (2002, pág. 396)	89
4.15	Tomado de Pérez, Sanjuán y Bermúdez (2002, pág. 586)	91
4.16	Tomado de Conejo y col. (2003, pág. 218)	92
4.17	Tomado de Docampo (2002, pág. 575)	94
4.18	Tomada de Chica y Lupiañez (2004, pág. 251)	96
4.19	Tomada de Santostefano, Quiroga y Rooney (2001, pág. 44)	99
4.20	Tomado de Guzmán y col. (2004, pág. 445)	103
4.21	Preparada a partir de la Tabla 3 de Martínez y col. (2000; pág. 366)	104
4.22	Tomado de Miñarro y col. (2000, pág. 650)	106
4.23	Tomado de Sierra y col. (2002, pág. 780)	110
4.24	Tomado de Alarcón y Blanca (2000, pág. 16)	116
4.25	Tomada de González y col. (2000, pág. 552)	124
4.26	Tomada de Sánchez y col. (2002, pág. 705)	128
4.27	Tomada de Alvarado y Santisteban (2004, pág. 279)	130
4.28	Preparada a partir de la Tabla 1 de Marcos y Redondo (2005; pág. 335))	134
4.29	Tomada de Tornay y Milán (2001, pág. 114)	136
4.30	Tomado de Rodríguez y Alonso (2004, pág. 211)	139
4.31	Tomado de Betancor y col. (2003, pág. 410)	142
5.1	Diseño 2x3: representación gráfica de las medias de las casillas	163
5.2	Diseño 2x3. Representación gráfica de las medias	166
5.3	Comparaciones entre efectos simples de interacción	185
5.4	Interacción entre <i>tratamiento</i> y <i>hacinamiento</i> en cada <i>momento del día</i>	198
5.5	Interacción entre <i>tratamiento</i> y <i>momento</i> en cada nivel de <i>hacinamiento</i>	199